Gravitational Form Factors through the application of novel computational techniques

PI: Alexandre Camsonne,
CO-PI: Volker Burkert, David Richards
Collaborators: **Melany Higuera-Angulo,** Pierre Chatagnon, Herve Dutrieux, Christian Chamness

Mechanical properties of hadrons: Structure, dynamics, visualization
ECT* Trento, Italy
March 3th, 2025

Jefferson Lab

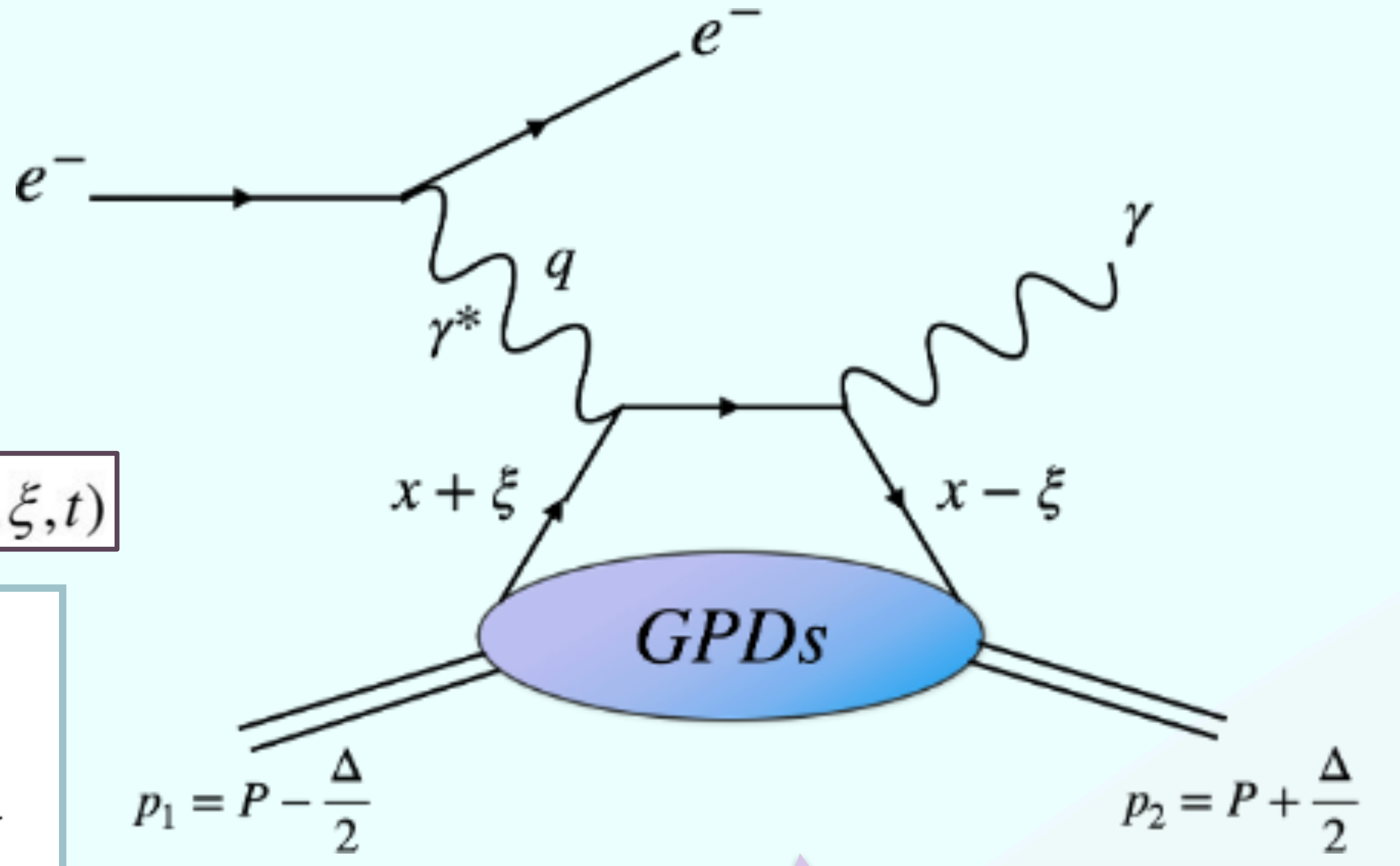
Outline

1. Motivation and scope of the project:
   - Mechanical properties of the proton.
   - Accessing GFFs through 3D distributions: Compton form factors.
2. Framework:
   - Extraction of Gravitational Form Factors using Neural Networks, from experimental data with insights from LQCD.
   - Technical difficulties of using NNets.
3. State of the project.
4. Deliverables and future directions.

Motivation and scope of the project:

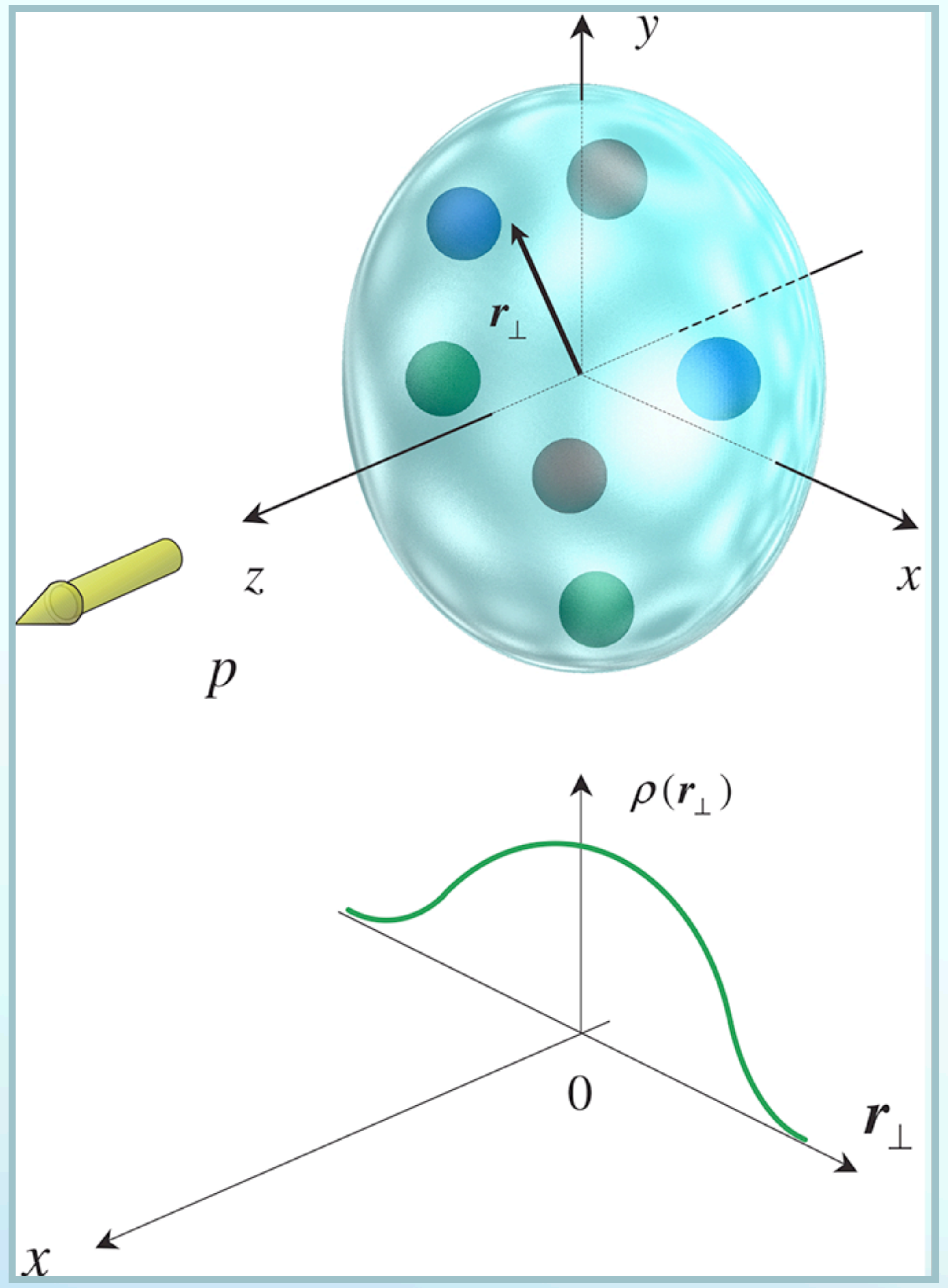
- Mechanical properties of the proton.
- Extraction of Gravitational Form Factors using Neural Networks, from experimental data with insights from LQCD.

3D distributions: Generalized Parton Distributions

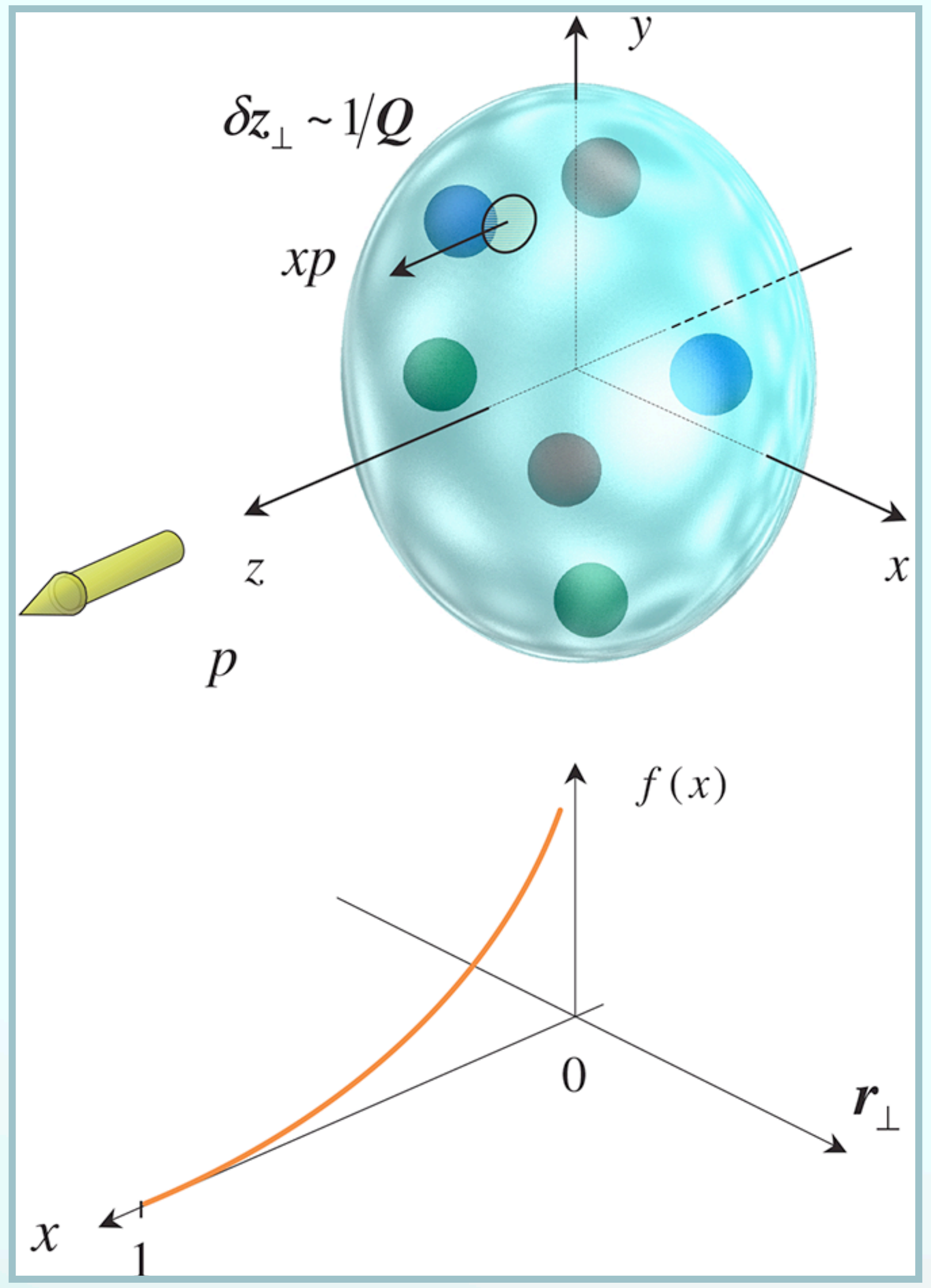


- GPDs provide a kaleidoscopic view of the 3D spatial structure of hadrons.

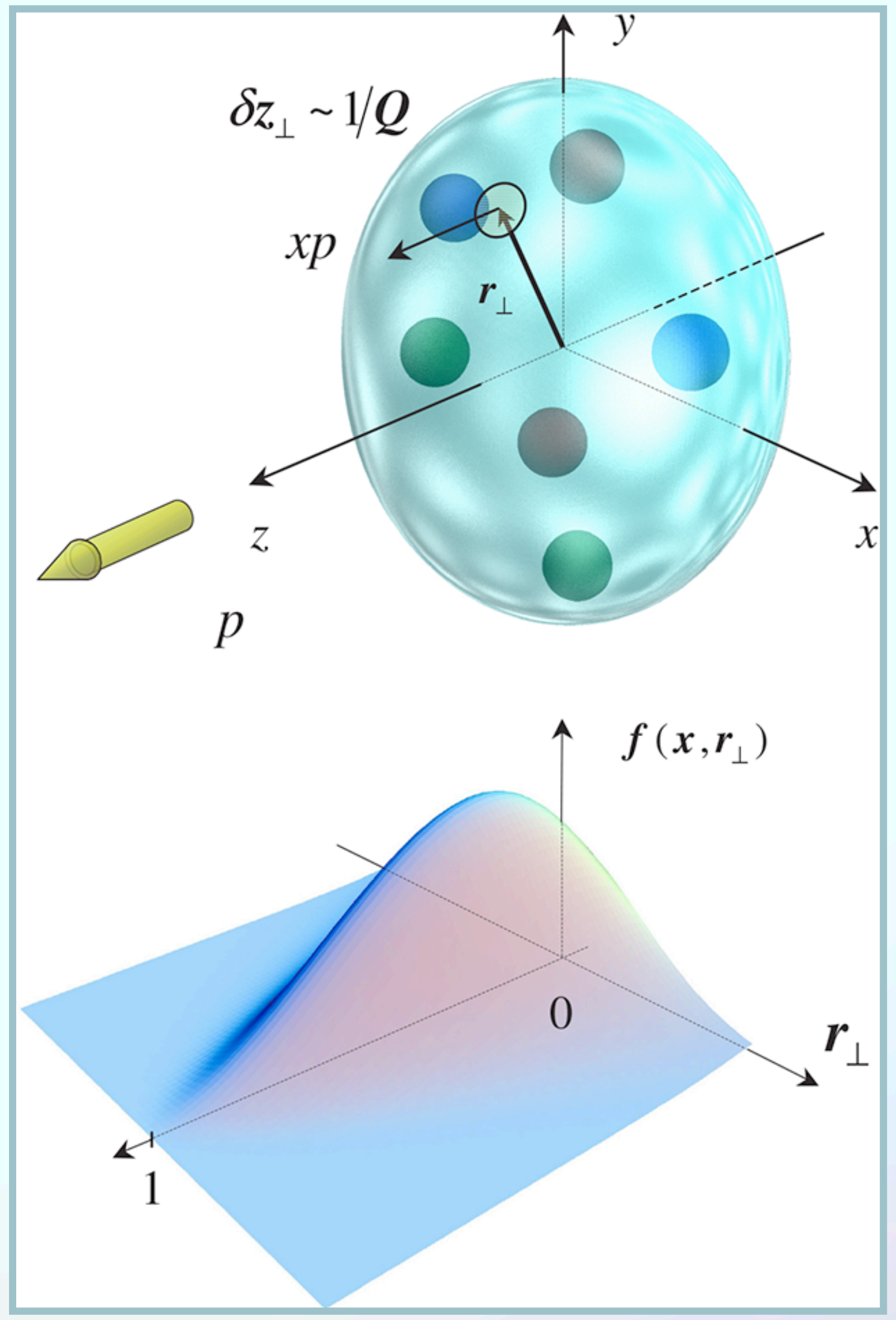
$H(x,\xi,t), E(x,\xi,t), \tilde{H}(x,\xi,t), \tilde{E}(x,\xi,t)$



Transverse charge and current density:
Form Factors (FFs)



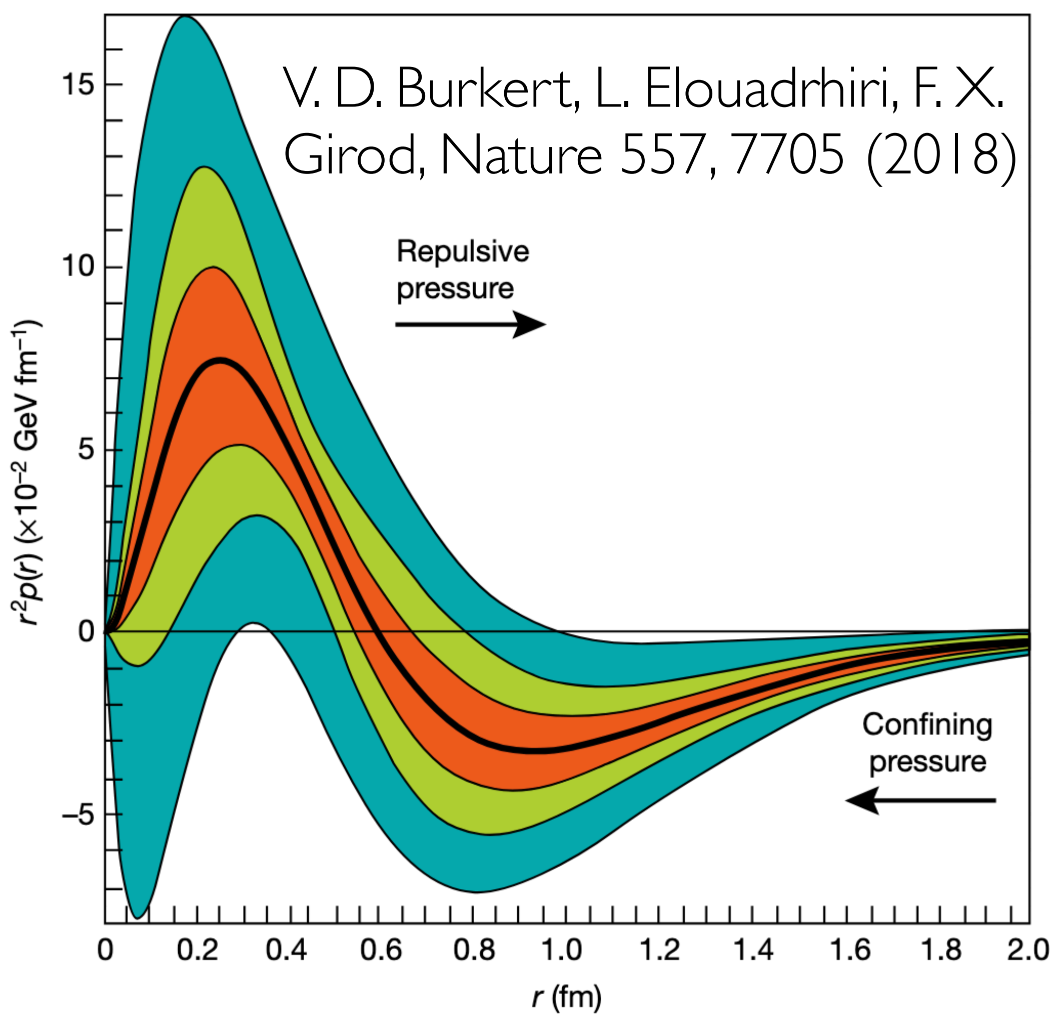
DIS structure functions:
Parton Distribution Functions (PDFs)

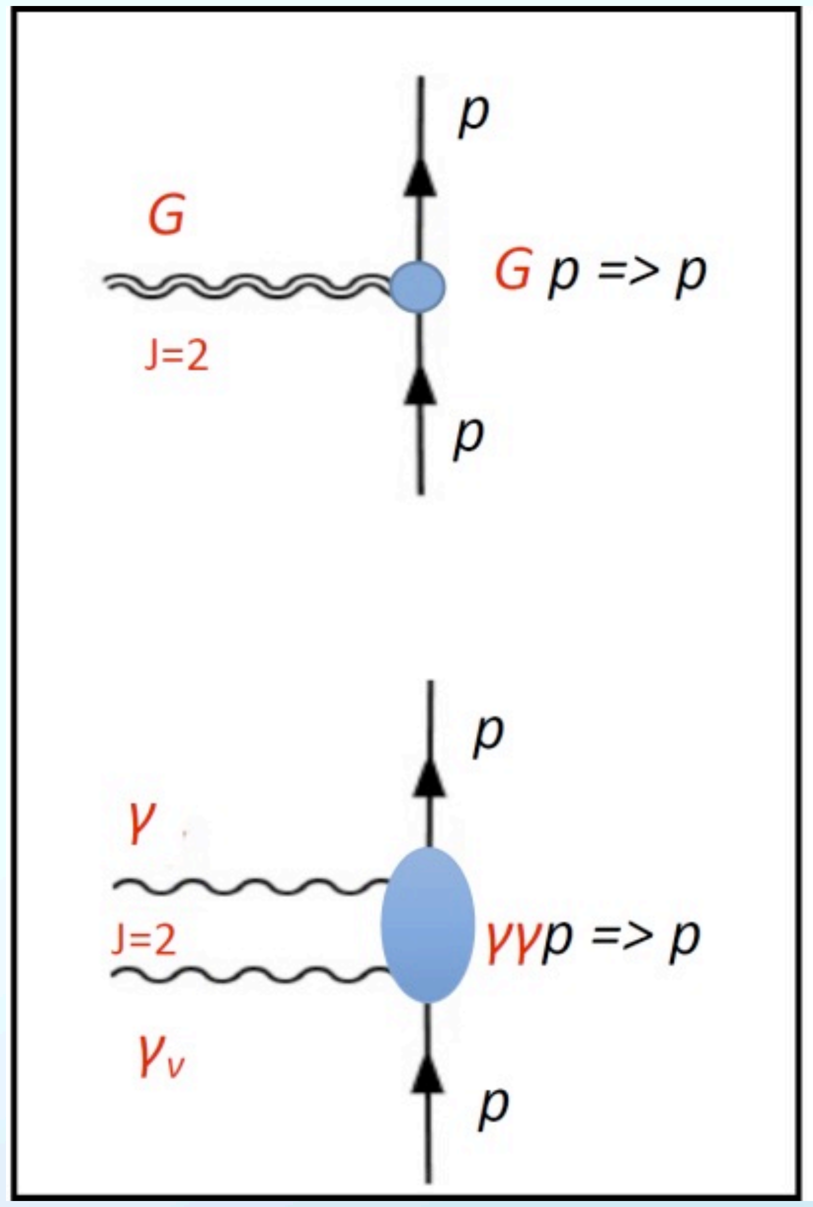


3D image in transverse space and longitudinal momentum:
Generalized Parton Distributions (GPDs)

Exclusive process:
Deeply Virtual Compton Scattering (DVCS)
$(e^- p \rightarrow e^- p \gamma)$

Mechanical properties of the proton





What are these mechanical properties?

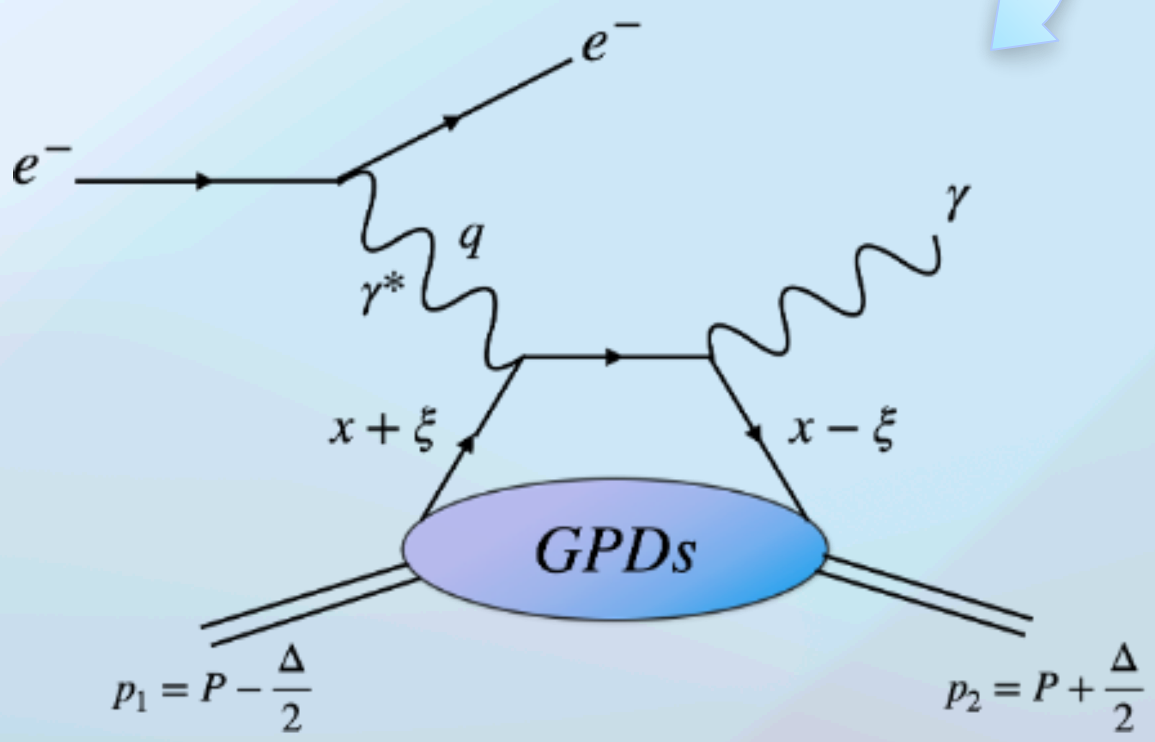
- Internal distribution of forces and pressure.
- Distribution of mass and energy.
- Distribution of angular momentum.
- Physical size as given by the mechanical radius.

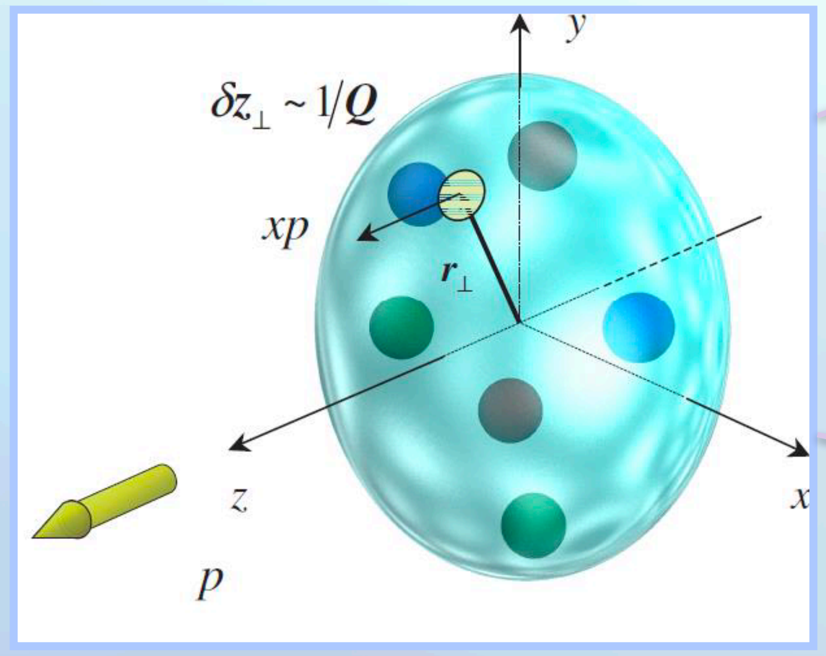
There is no direct way to measure the gravitational form factors (GFFs).

However, they may be probed indirectly in various exclusive processes like DVCS that simulate the graviton-proton interaction.

The 2γ field couples to the EMT as gravity does with a strength many orders of magnitude greater.

The leading contribution to DVCS is described in terms of four generalized parton distributions (GPDs). Two of them give access to the quark GFFs. Therefore it also let to the insight that GPDs can be related to the mechanical properties of nucleons.



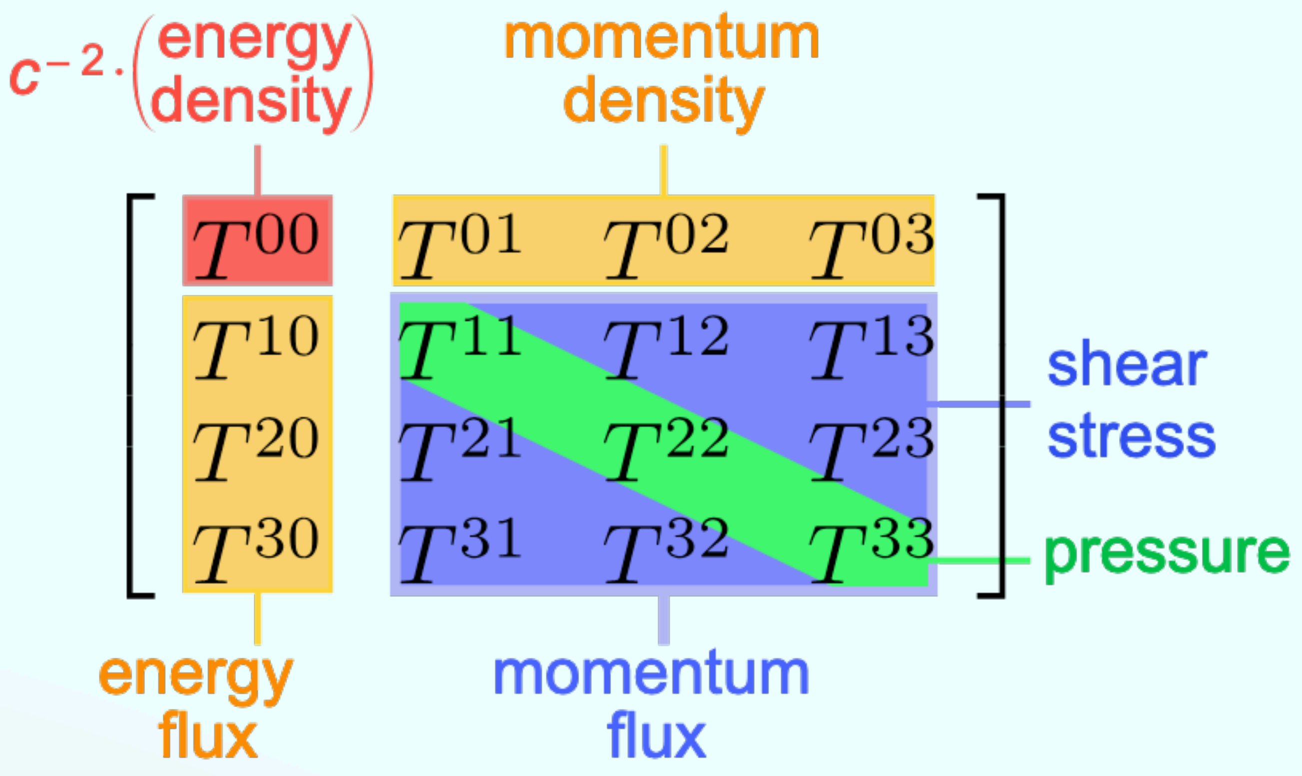


3D quark distributions

$$\int_{-1}^{1} \mathrm{d}x\, x H_q(x,\xi,t) = A_q(t) + \xi^2 D_q(t),$$
$$\int_{-1}^{1} \mathrm{d}x\, x E_q(x,\xi,t) = B_q(t) - \xi^2 D_q(t),$$

Mechanical properties of the proton

The Matrix element of the QCD Energy-Momentum tensor encodes key information of the proton including the mass and spin, as the distributions of energy, angular momentum, and various mechanical properties such as, e.g., internal forces inside the system.



$$\langle p',\vec{s}'|T_a^{\mu\nu}|p,\vec{s}\rangle = \bar{u}(p',\vec{s}')\Bigg[A_a(t)\frac{P^\mu P^\nu}{M_N} + D_a(t)\frac{\Delta^\mu\Delta^\nu - g^{\mu\nu}\Delta^2}{4M_N} + \bar{C}_a(t)M_N\,g^{\mu\nu} + J_a(t)\frac{P^{\{\mu}i\sigma^{\nu\}\lambda}\Delta_\lambda}{M_N} - S_a(t)\frac{P^{[\mu}i\sigma^{\nu]\lambda}\Delta_\lambda}{M_N}\Bigg]u(p,\vec{s})$$

Momentum — Pressure — Anisotropy — Total angular momentum — Intrinsic angular

Symmetric kinematical variables:

$P = (p'+p)/2$ and $\Delta = p'-p$

Free Dirac spinor: $u(p,\vec{s})$

GFFs:

$A_a(t)$, $D_a(t)$, $\bar{C}_a(t)$, and $J_a(t)$

On top of restricting the number of GFFs, Poincaré symmetry imposes additional constraints:

$$\begin{aligned} A(0) &= \sum_q A_q(0) + A_G(0) = 1, \\ J(0) &= \sum_q J_q(0) + J_G(0) = \tfrac{1}{2}, \\ \tfrac{1}{2}\Delta\Sigma &= \sum_q S_q(0), \\ \bar{C}(t) &= \sum_q \bar{C}_q(t) + \bar{C}_G(t) = 0, \end{aligned}$$

V. D. Burkert, et. al. Rev. Mod. Phys. 95, 041002

D(t) is unknown. It value is not fixed by spacetime symmetries and must be determined from experiment.

$$D(t) = ? \quad \Rightarrow \quad \Delta(t) \propto D(t) \propto \int d^3\mathbf{r}\; p(r)\frac{j_0(r\sqrt{-t})}{t}$$

The D-term is not related to "external properties" of a particle like mass and spin, but to the stress tensor and internal force.

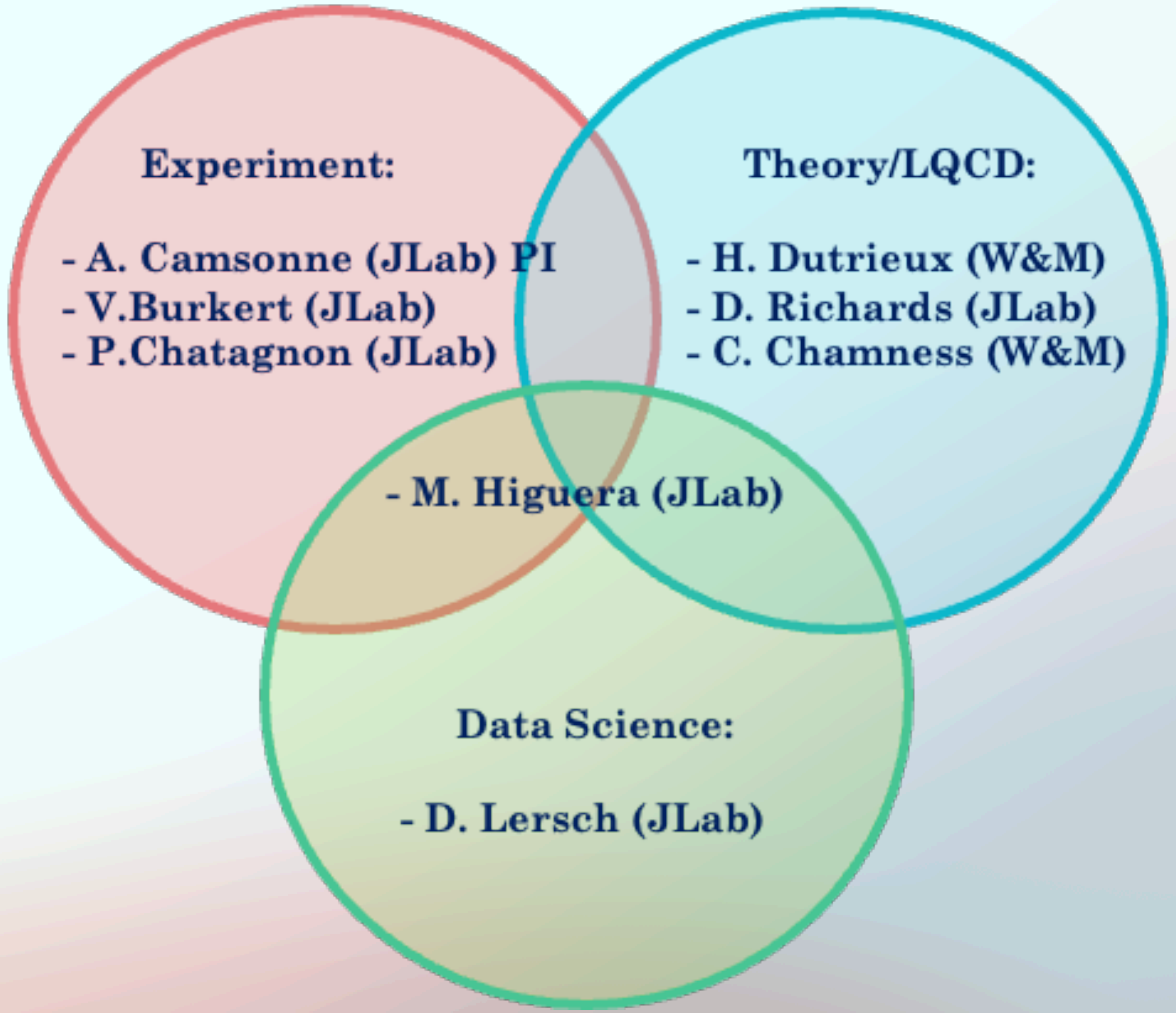
Our goal:

- Make a global analysis using NN

  ➡ Extract GFFs

- Use state-of-the-art ML libraries: Pytorch.

- Handle experimental and LQCD data.



Accessing GFFs through 3D distributions: Compton form factors

Complications: The actual observables in DVCS are the Compton form factors (CFFs) which are expressed by means of factorization formulae in terms of complex-valued convolution integrals:

$$\mathrm{Re}\mathcal{H}(\xi,t) + i\,\mathrm{Im}\mathcal{H}(\xi,t) = \sum_q e_q^2 \int_{-1}^{1} dx \left[\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon}\right] H_q(x,\xi,t)$$

Where $\xi = x_B/(2 - x_B)$, x_B is the Bjorken scaling variable, and t is the squared momentum transfer to the nucleon.

- This convolution integral cannot be inverted in a model-independent way to yield GPDs.

Thanks to the DVCS CFFs it is possible to obtain the fixed-t dispersion relation (DR) where the subtraction constant is related to the so-called D-term.

$$\mathrm{Re}\mathcal{H}(\xi,t) = \mathcal{C}_{\mathcal{H}}(t) + \frac{1}{\pi}\,\mathrm{P.V.}\int_0^1 d\xi' \left[\frac{1}{\xi - \xi'} - \frac{1}{\xi + \xi'}\right] \mathrm{Im}\mathcal{H}(\xi',t)$$

Subtraction constant

The CFFs are linked to measurable quantities, such as differential cross sections and beam, charge, and target polarization asymmetries.

Accessing GFFs

Considering some assumptions like:

1. At LO, gluons don't contribute to DVCS much, so one can assume that the subtraction constant is dominantly coming from quarks, but for the moment we neglect also strange and heavier quarks.

2. To talk about "quark pressure" one needs to neglect energy-momentum flow from quarks to gluons described by EMT form factor $\bar{C}(t)$.

2. In the asymptotic limit $d_n^q(t)$ vanish for n>1. Where $d_1^q(t)$ determine the symptomatic form of the GPDs.

3. Considerate the dominance of the flavor singlet combination.

$$\mathcal{C}_{\mathcal{H}}(t) = 2\sum_q e_q^2 \int_{-1}^{1} \mathrm{d}z \, \frac{D^q_{\text{term}}(z,t)}{1-z},$$

$$d_1^u = d_1^d = d_1^Q/2$$

$$D^q_{\text{term}}(z,t) = (1-z^2) \sum_{\text{odd } n} d_n^q(t)\, C_n^{3/2}(z)$$

$$D_q(t) = \frac{4}{5} d_1^q(t) = \int_{-1}^{1} \mathrm{d}z\, z\, D^q_{\text{term}}(z,t)$$

$$\Delta(t) = 4\left(\frac{4}{9} d_1^u(t) + \frac{1}{9} d_1^d(t)\right)$$

$$D^Q(t) = \frac{18}{25}\Delta(t)\,.$$

To date, a more conservative extraction of the subtraction constant from the currently available experimental data remains compatible with zero within large uncertainties.

By utilizing this DR, along with other phenomenological assumptions, we can extract the CFFs and GFFs from experimental data and enhance the results with lattice calculations!

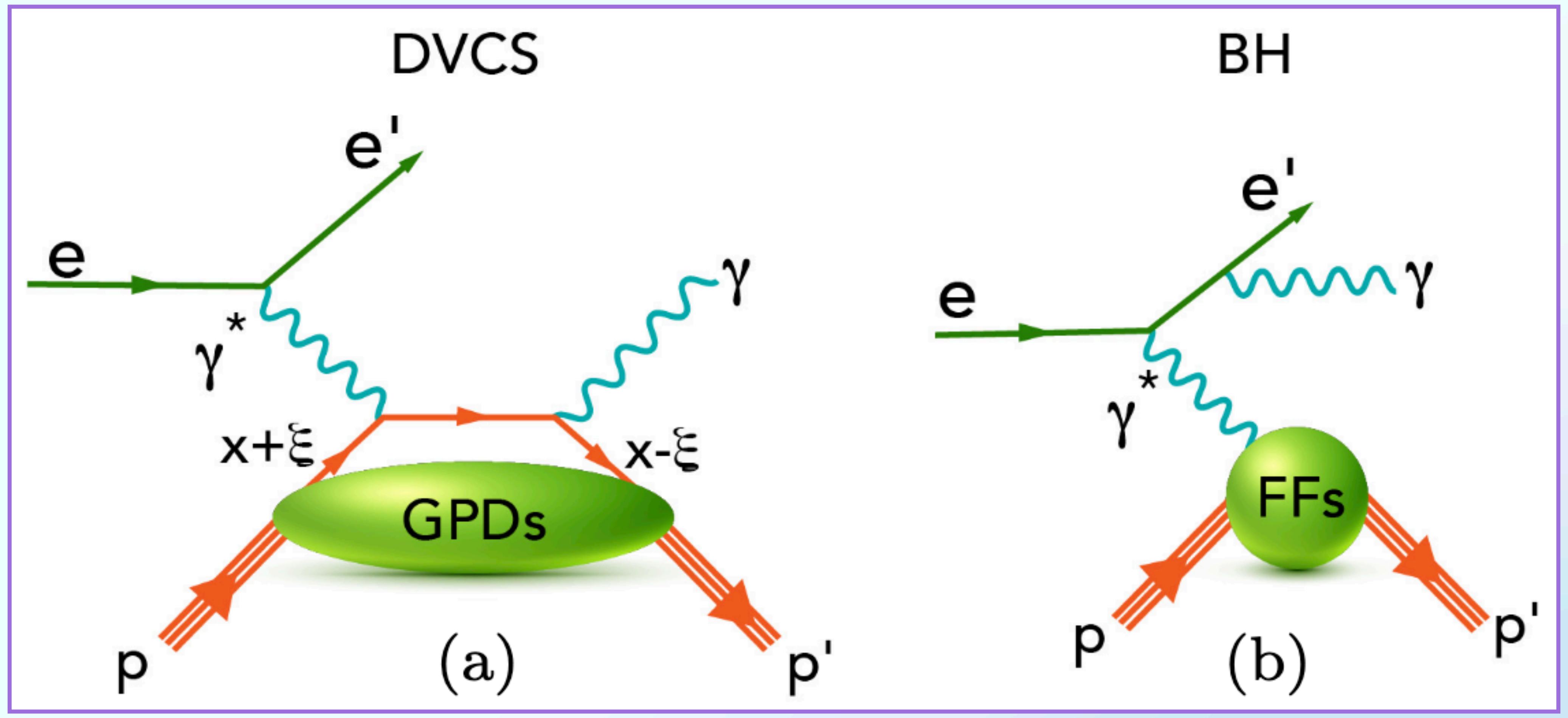
Matching theory with experiment

Measurements of DVCS are mostly realized via the process of leptoproduction of a real photon, where also an interference with the Bethe-Heitler radiation occurs

The BH process gives access to elastic form factors, it does not see the 3D structure of hadrons.

$$\frac{d^5\sigma}{dx_B dQ^2 d|t| d\phi d\phi_S} = \frac{\alpha^3 x_B}{16\pi^2 Q^4 \sqrt{1+\epsilon^2}} |\mathcal{T}|^2 ,$$

$$|\mathcal{T}|^2 = |\mathcal{T}_{\mathrm{BH}} + \mathcal{T}_{\mathrm{DVCS}}|^2 = |\mathcal{T}_{\mathrm{BH}}|^2 + |\mathcal{T}_{\mathrm{DVCS}}|^2 + \mathcal{I} .$$



The DVCS amplitude can be decomposed either in helicity amplitudes or, equivalently, in complex valued CFFs.

$$|\mathcal{T}_{\mathrm{DVCS}}|^2 = \frac{2(2-2y+y^2)}{y^2 Q^2 (2-x_{\mathrm{B}})^2} \Big[4(1-x_{\mathrm{B}}) \left(|\mathcal{H}|^2 + |\widetilde{\mathcal{H}}|^2 \right) - \left(x_{\mathrm{B}}^2 + (2-x_{\mathrm{B}})^2 \frac{\Delta^2}{4M^2} \right) |\mathcal{E}|^2 - x_{\mathrm{B}}^2 \left(\mathcal{H}\mathcal{E}^* + \mathcal{E}\mathcal{H}^* + \widetilde{\mathcal{H}}\widetilde{\mathcal{E}}^* + \widetilde{\mathcal{E}}\widetilde{\mathcal{H}}^* \right) - x_{\mathrm{B}}^2 \frac{\Delta^2}{4M^2} |\widetilde{\mathcal{E}}|^2 \Big], \quad y = \frac{Q^2}{xs}.$$

Thursday's talk by Marija Cuic

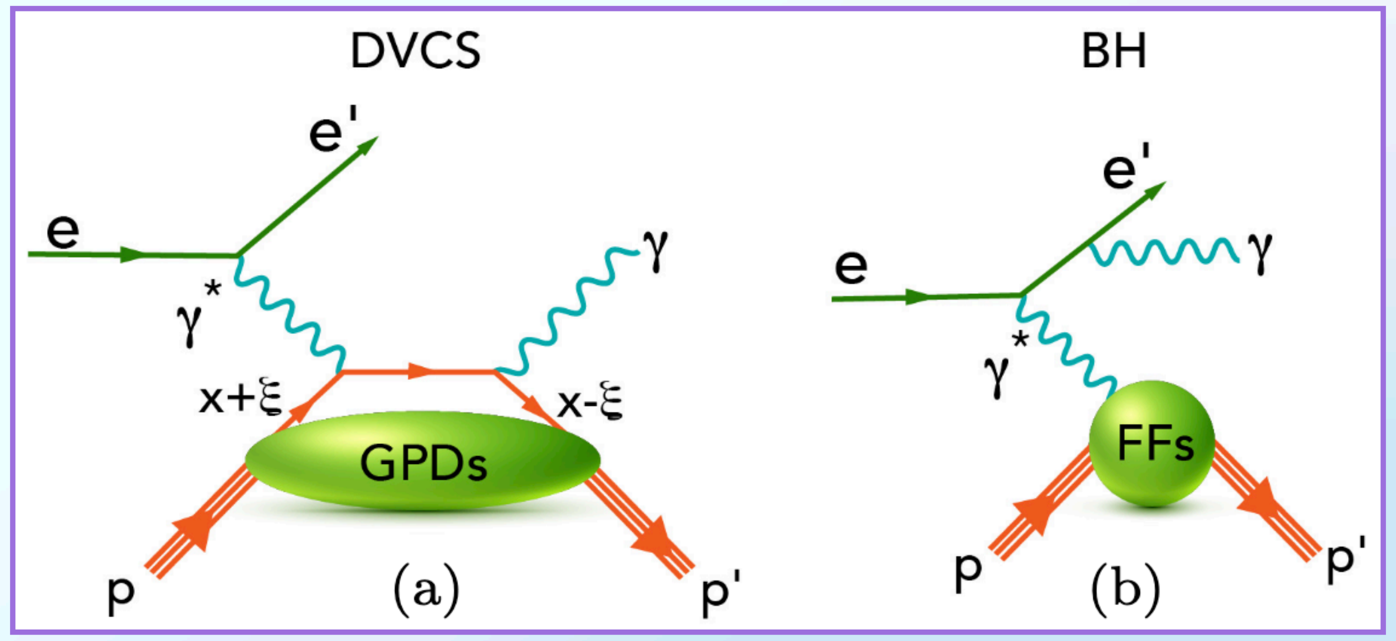
Jefferson Lab

Experimental observables

For observables like asymmetries, it is possible to still access more or less linear CFFs using the dominance of the Bethe-Heitler (BH) process in the denominator of the DVCS amplitude.

For example, the first sine harmonic of the beam spin asymmetry (BSA), as measured e.g. in Jefferson Lab is defined as:

$$A_{LU}^{-,\sin\phi} \equiv \frac{1}{\pi}\int_{-\pi}^{\pi} d\phi\, \sin\phi\, A_{LU}^{-}(\phi)$$



This harmonics are then proportional to linear combinations of CFFs:

$$A_{LU}^{-,\sin\phi} \propto \mathfrak{Im}\left(F_1\mathcal{H} - \frac{t}{4M^2}F_2\mathcal{E} + \frac{x_B}{2}(F_1+F_2)\tilde{\mathcal{H}}\right)$$

If cross-sections are measured, using weighted Fourier integrals can help to cancel the strongly oscillating Bethe-Heitler propagators and interference terms improving the convergence of harmonic terms

- **Framework:** Extraction of Gravitational Form Factors using Neural Networks, from experimental data.
- Parallel projects: Database.
- Technical difficulties of using NNets.

Framework

We are building upon existing tools: Gepard

- Python software framework dedicated to the study of GPDs.
- K. Kumerički, D. Müller, K. Passek-Kumerički, Towards a fitting procedure for deeply virtual Compton scattering at next-to-leading order and beyond, Nuclear Physics B, Volume 794, Issues 1–2, 2008

To obtain a more reliable estimate of uncertainties with reduced model dependence, we decided to utilize Artificial Neural Networks to parameterize the CFFs.

- Use State-of-the-art standard ML libraries. PyTorch
- A tool to fit all existing data using NN to extract GFFs.

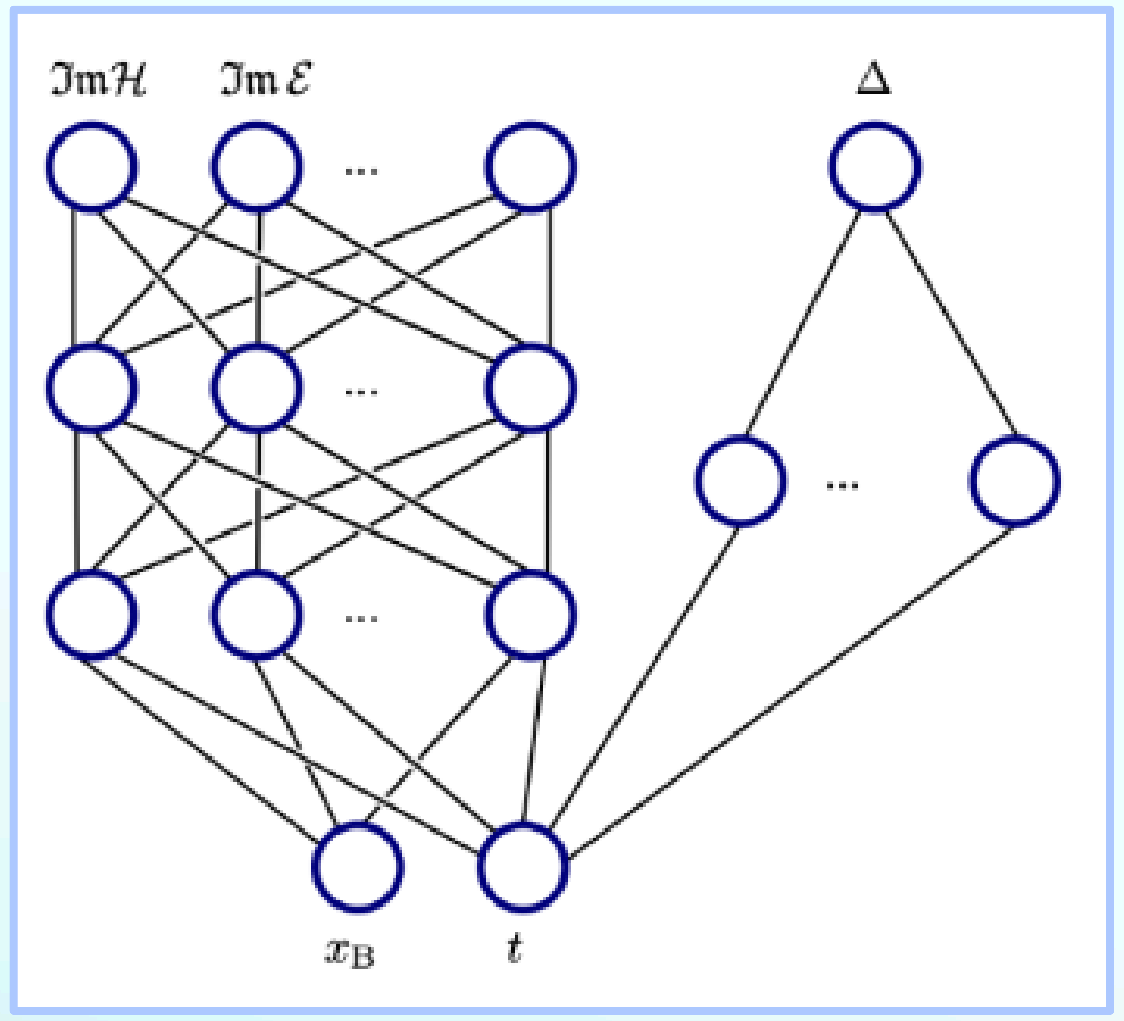
Framework

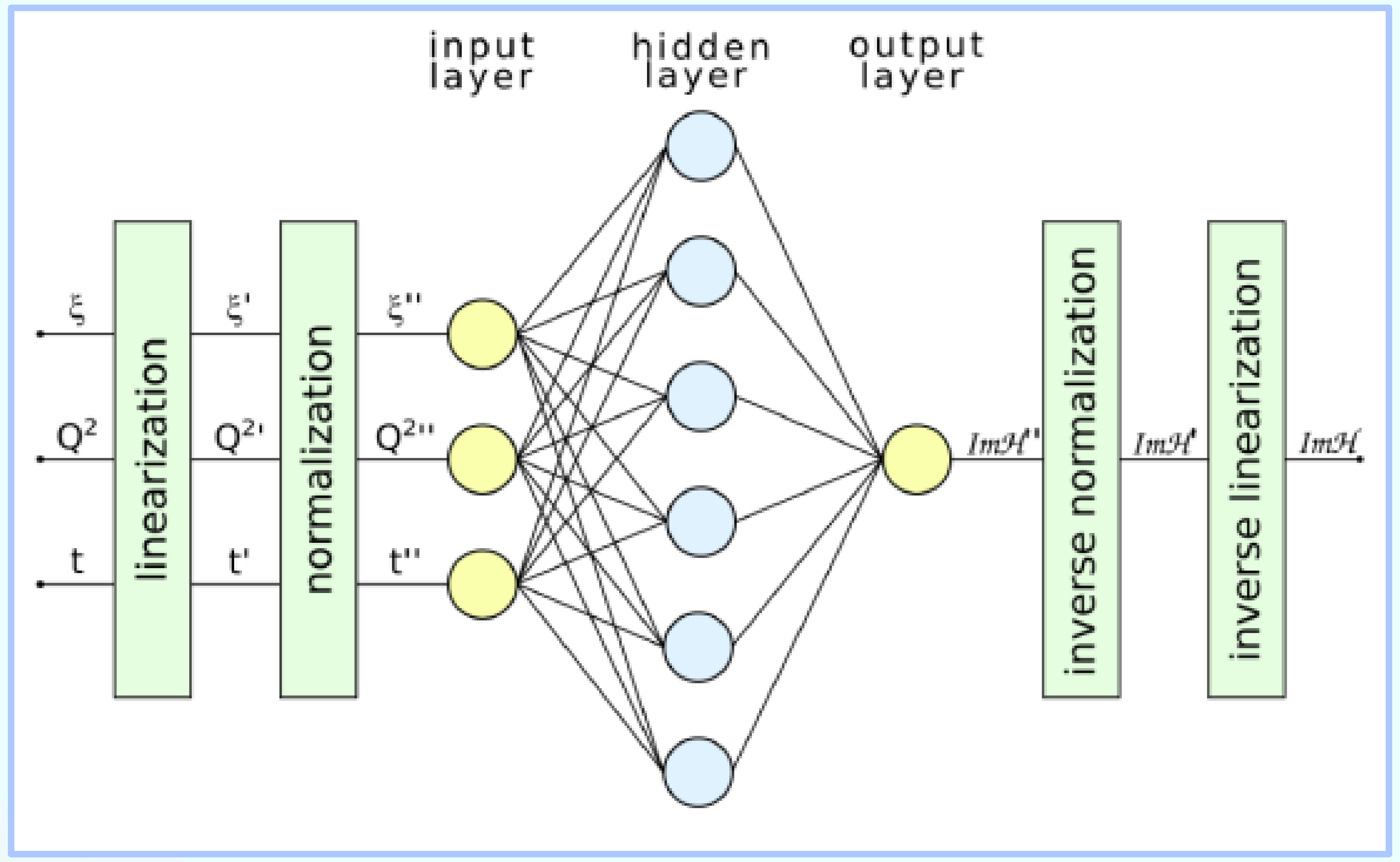
Using Gepard

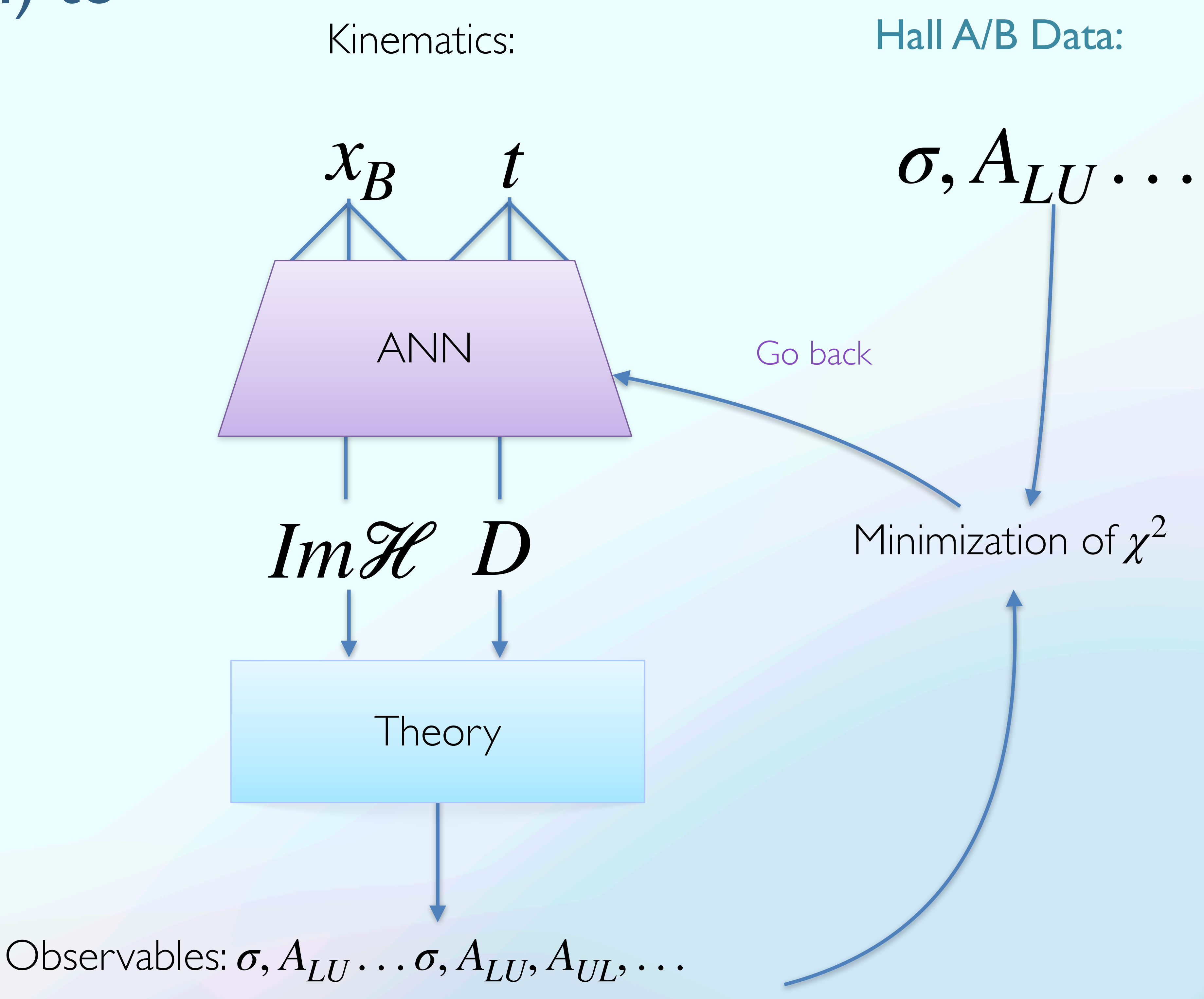
With opengpd / gpddatabase

Artificial Neural Networks (ANN)

A machine learning (ML) algorithm that uses artificial intelligence (AI) to learn how to process data by analyzing training examples.







$$\mathrm{Re}\mathcal{H}(\xi,t) = \mathcal{C}_{\mathcal{H}}(t) + \frac{1}{\pi}\,\mathrm{P.V.}\int_0^1 d\xi' \left[\frac{1}{\xi-\xi'} - \frac{1}{\xi+\xi'}\right] \mathrm{Im}\mathcal{H}(\xi',t)$$

It's essentially a least-squares fit of a complex, multi-parameter function, implying no theoretical bias.

- Fits are performed in harmonic space, therefore harmonic analysis (Fourier transform) of the data was previously necessary.

Database

opengpd / gpddatabase

Currently, there was no unified common database storing all available data which is necessary to reduce the uncertainties.

We developed a publicly accessible database format capable of storing current and future data as well as pseudodata.

- Stores data and pseudodata coming from various elastic and exclusive measurements and lattice-QCD results.
- Interfaces are provided in the main programming languages: Python and C++.
- Interface with GEPARD and Parsons.
- Accessible by anyone in the world using Github.

Database for studying Generalised Parton Distributions (GPDs)

This page:
about
installation
data format
basic usage
available data
adding new data

license
acknowledgements
contact

View GitHub repository:
link

View PyPI repository:
link

View reference article:
arXiv inspire

Available datasets

uuid	collaboration	reference	type	pseudo	observables	comment
EqbtDRkv	HallA	https://arxiv.org/pdf/1703.09442.pdf	DVCS	False	CrossSectionDifferenceLU	None
PusMstKs	CLAS	https://arxiv.org/pdf/2211.11274	DVCS	False	ALL	None
vGAKAf7P	CLAS	https://arxiv.org/pdf/1501.07052.pdf	DVCS	False	ALU AUL ALL	None
75ueQoQw	COMPASS	https://arxiv.org/abs/1802.02739	DVCS	False	CrossSectionUUVirtualPhotoProduction TSlope	None
ob8hLTm2	CLAS	https://arxiv.org/pdf/1810.02110	DVCS	False	CrossSectionUU	Differential CrossSection for exclusive DVCS electroproduction
3gYp9R4P	HERMES	https://arxiv.org/pdf/hep-ex/0605108.pdf	DVCS	False	AcCos1Phi	None
bmTzHHvg	HallA	https://arxiv.org/pdf/2109.02076	DVCS	False	CrossSectionUU	Beam-helicity-independent cross section
RQncbKtk	CLAS	https://arxiv.org/pdf/1504.02009v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None
AtY8o7Ej	HallA	https://arxiv.org/pdf/1504.05453v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None

In collaboration with P. Sznajder and C. Mezrag

arxiv.org/abs/2503.18152

https://opengpd.github.io/gpddatabase/index.html

Database

opengpd / gpddatabase

Currently, there was no unified common database storing all available data which is necessary to reduce the uncertainties.

We developed a publicly accessible database format capable of storing current and future data as well as pseudodata.

- Get the list of available data files.
- Load a given data file
- Get metadata.
- Accessible by anyone in the world using Github.

Database for studying Generalised Parton Distributions (GPDs)

This page:
about
installation
data format
basic usage
available data
adding new data

license
acknowledgements
contact

View GitHub repository:
link

View PyPI repository:
link

View reference article:
arXiv inspire

Available datasets

uuid	collaboration	reference	type	pseudo	observables	comment
EqbtDRkv	HallA	https://arxiv.org/pdf/1703.09442.pdf	DVCS	False	CrossSectionDifferenceLU	None
PusMstKs	CLAS	https://arxiv.org/pdf/2211.11274	DVCS	False	ALL	None
vGAKAf7P	CLAS	https://arxiv.org/pdf/1501.07052.pdf	DVCS	False	ALU AUL ALL	None
75ueQoQw	COMPASS	https://arxiv.org/abs/1802.02739	DVCS	False	CrossSectionUUVirtualPhotoProduction TSlope	None
ob8hLTm2	CLAS	https://arxiv.org/pdf/1810.02110	DVCS	False	CrossSectionUU	Differential CrossSection for exclusive DVCS electroproduction
3gYp9R4P	HERMES	https://arxiv.org/pdf/hep-ex/0605108.pdf	DVCS	False	AcCos1Phi	None
bmTzHHvg	HallA	https://arxiv.org/pdf/2109.02076	DVCS	False	CrossSectionUU	Beam-helicity-independent cross section
RQncbKtk	CLAS	https://arxiv.org/pdf/1504.02009v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None
AtY8o7Ej	HallA	https://arxiv.org/pdf/1504.05453v1.pdf	DVCS	False	CrossSectionUU CrossSectionDifferenceLU	None

In collaboration with P. Sznajder and C. Mezrag

arxiv.org/abs/2503.18152

https://opengpd.github.io/gpddatabase/index.html

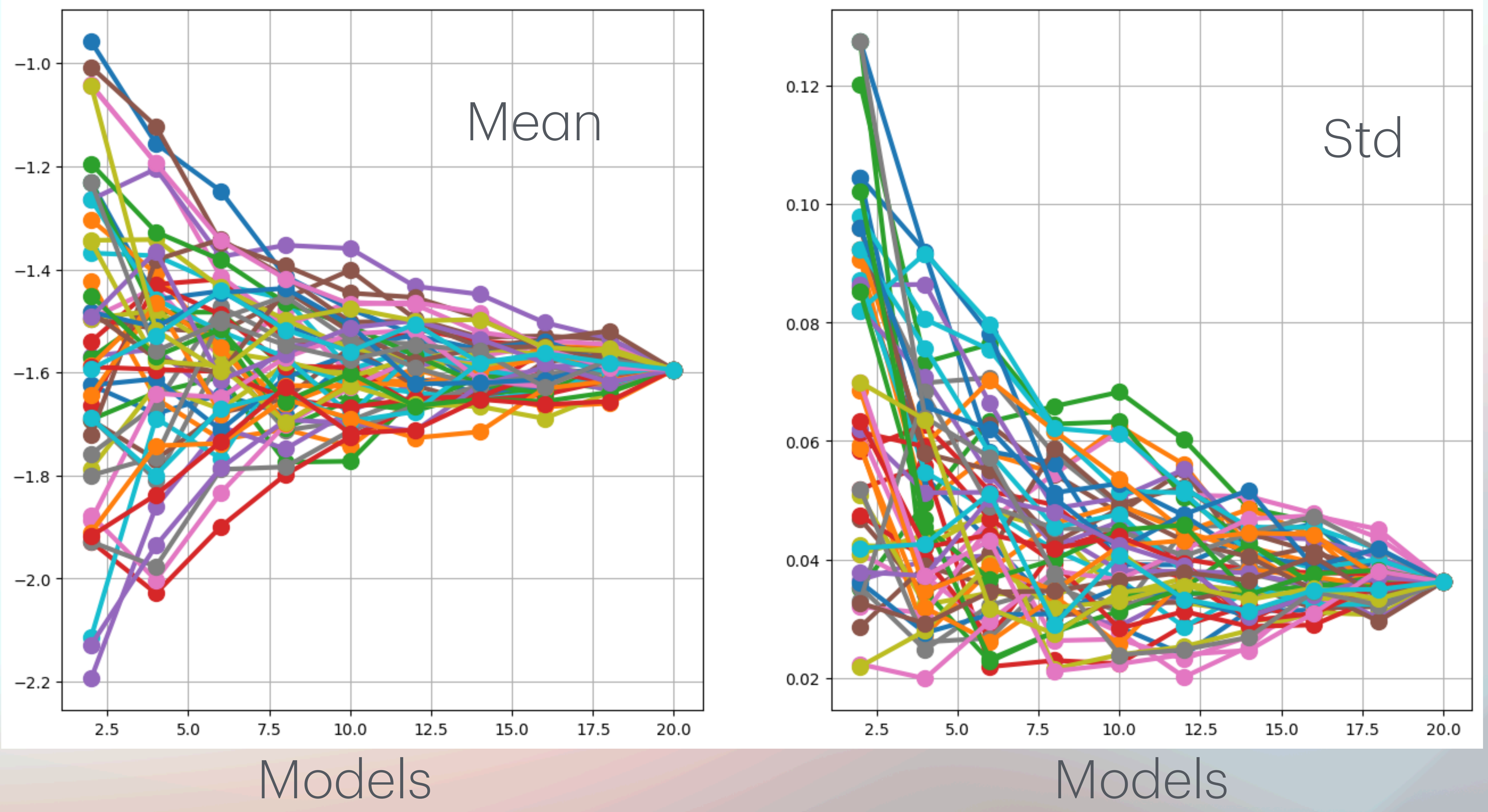
Jefferson Lab

Technical difficulties of using NNets

To solve instability in results we use ensemble analysis by averaging over multiple learned representations.

- This shows how as the number of models increases, the mean residual prediction tends to stabilize.

- The standard deviation plot shows how the uncertainty in our predictions increases as the ensemble shrinks



A larger ensemble tends to 'average out' individual model noise, improving the predictive stability results and uncertainty estimation of our model.

Jefferson Lab

State of the project

Experimental observables

Experimental data we have used:

- We started with a specific experiment.
- Describe 2007 and 2015 CLAS data to compare with previous results.
- Add the rest of the CLAS data at 6 GeV
- Add CLAS data at ~11 GeV.

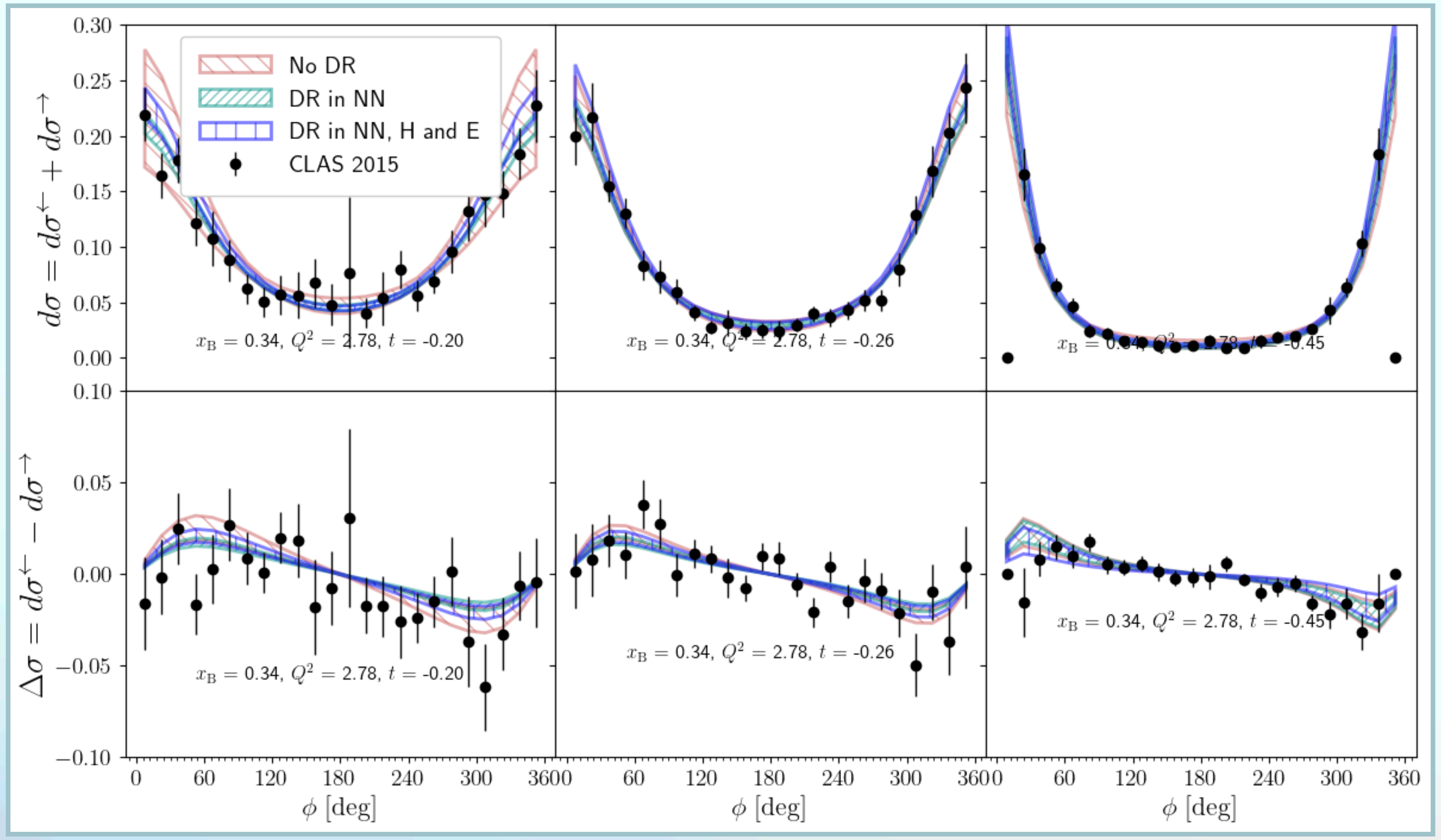
Collab	Year	Observable	l.b.e.	x_B	$Q^2[GeV^2]$	$\|t\|\,[GeV^2]$
CLAS	2001	$A_{LU}^{sin(\phi)}$	4.25	0.19	1.25	0.19
CLAS	2006	$A_{UL}^{sin(\phi)}$	5.7	0.2-0.4	1.82	0.15-0.44
CLAS	2007	A_{LU}	5.77	0.13-0.35	1.1-3	0.1-0.3
CLAS	2009	A_{LU}	4.8	0.12-0.48	1.0-2.8	0.09-1.8
CLAS	2015	$A_{LU}(\phi), A_{UL}(\phi), A_{LL}(\phi)$	5.93	0.18-0.4	1.6-3.2	0.1-0.45
CLAS	2015	$\sigma(\phi), \Delta\sigma(\phi)$	5.75	0.1-0.58	1-4.6	0.09-0.52
CLAS	2018	$\sigma(\phi)$	5.88	0.12-0.5	1.1-4	0.1-1.6
CLAS	2023	A_{LL}	10.2	0.09-0.45	1.3-6	0.1-2.2
CLAS	2023	A_{LL}	10.6	0.09-0.62	1.1-7.2	0.1-2.8

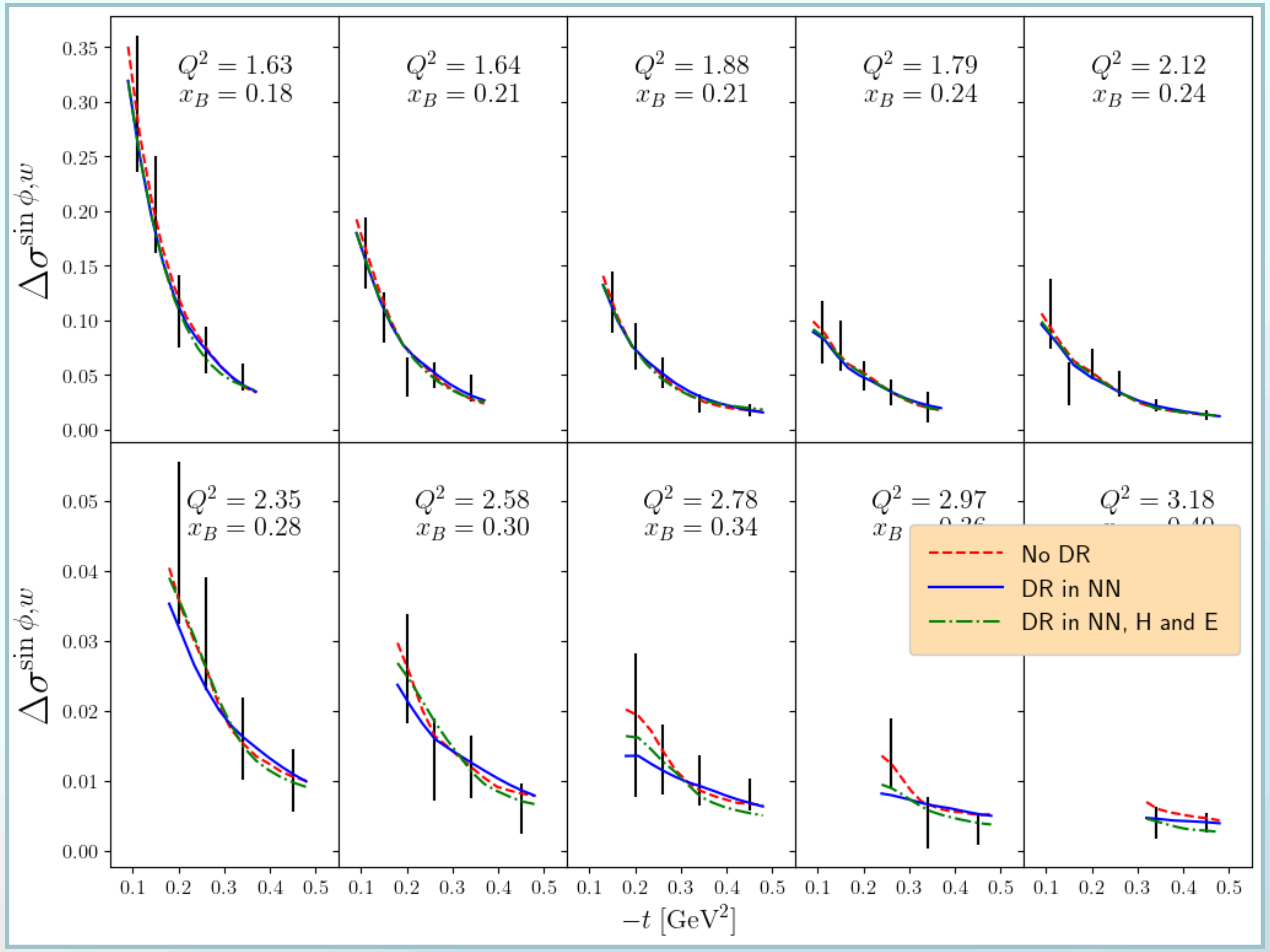
Lepton-beam-energy (l.b.e.)

We made cuts at t=0.51 GeV^2, and Q^2<1.5 GeV^2 or $-t/Q^2$<0.25

State of the project

Description of 2015 CLAS data through the NN



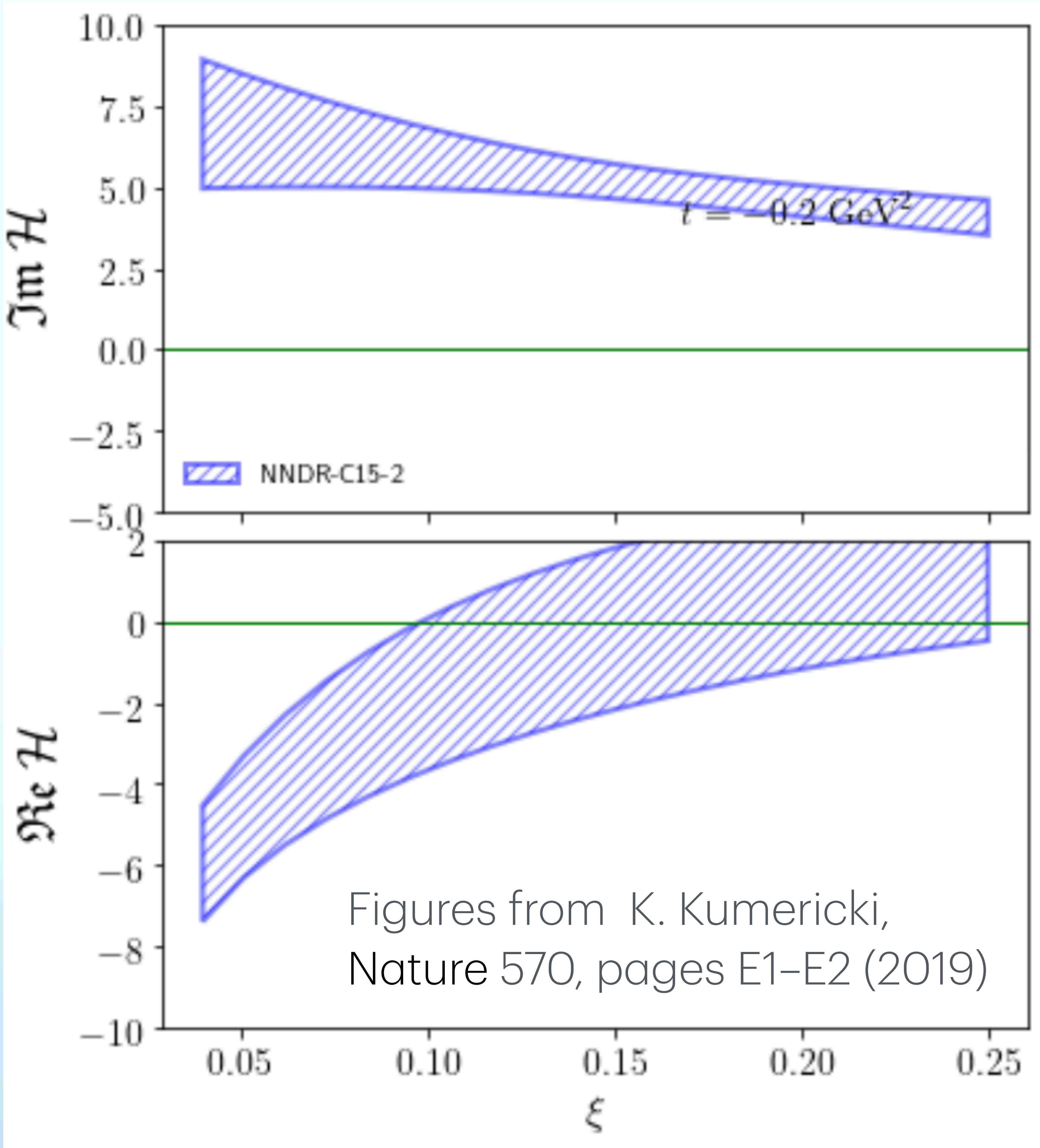


(CLAS Collaboration) Phys. Rev. Lett. 115, 212003,2015

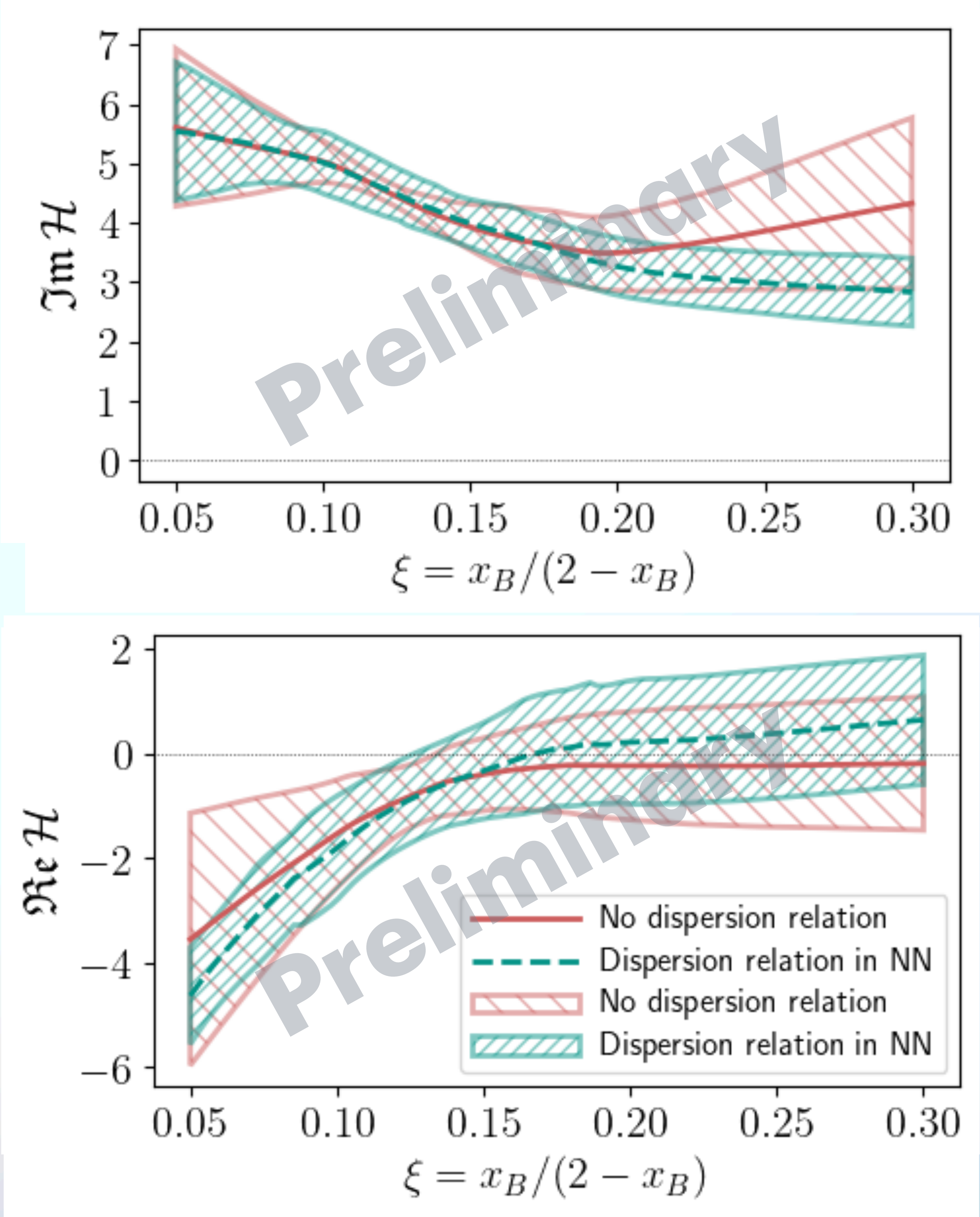
State of the project

- First, we used 2008 beam spin asymmetry and 2015 cross-section CLAS data to reproduce previous results.

CFFs using NN and dispersion relation



ImH and ReH CFFs



$\chi^2/dof = 0.31$

⇨ Qualitative good agreement with Kumericki's previous results.

⇨ • NN needs more constraints.

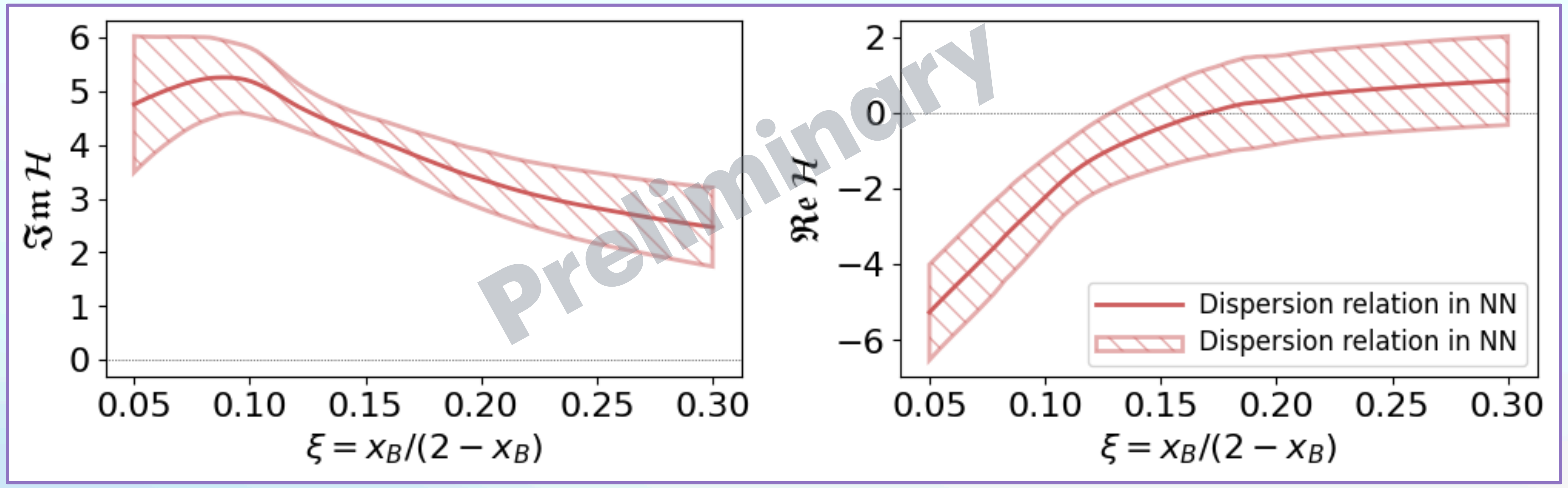
Jefferson Lab

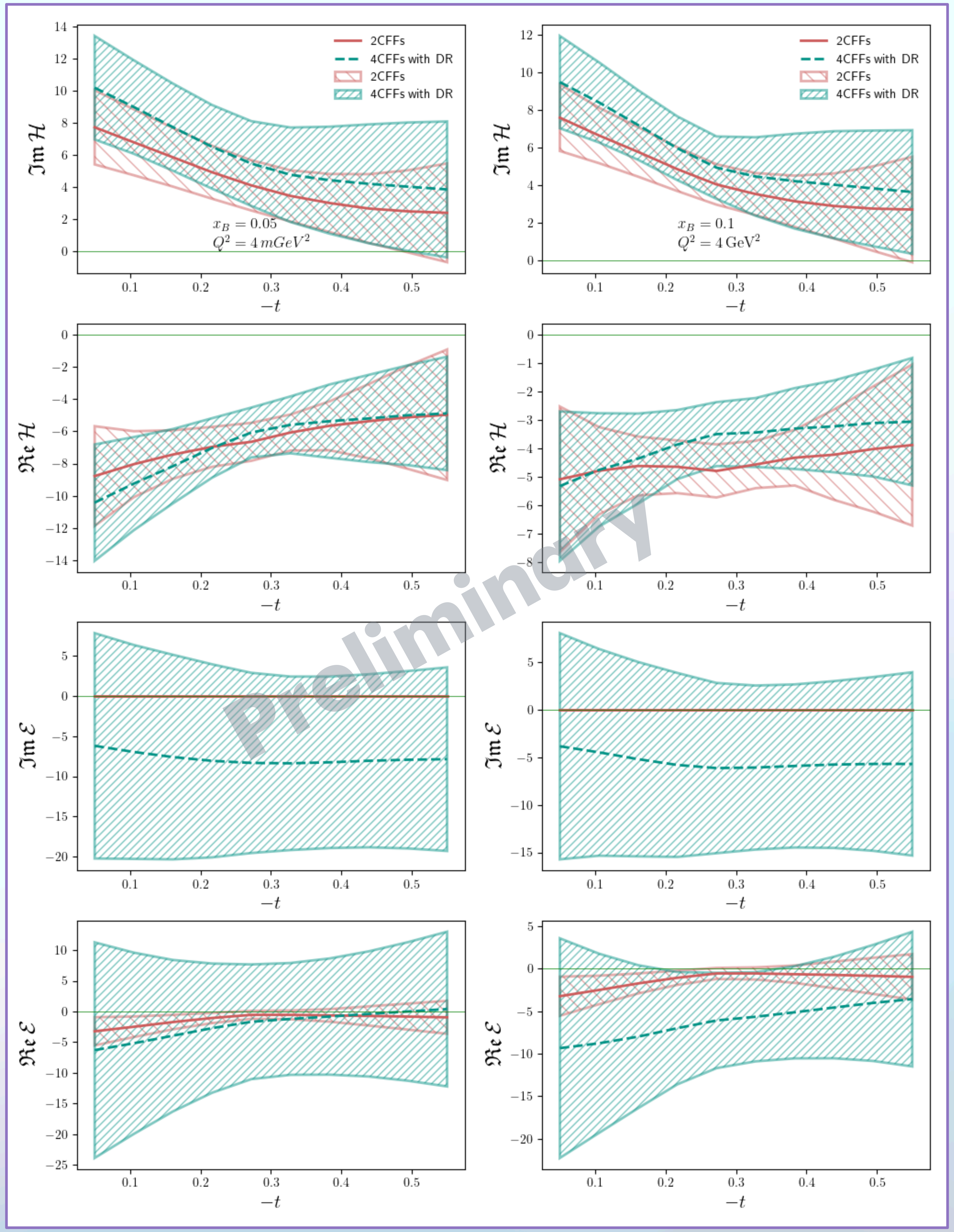
State of the project

Adding the rest of the previous CLAS data and new CLAS 2023 BSA data at 10.6 and 10.2 GeV (Phys. Rev. Lett. 130, 211902) to raise the range of kinematic restrictions

Extracting the harmonic spherics of the data, and analyzing the best neural nets parameters

4 CFFs





$\chi^2/dof = 0.38$

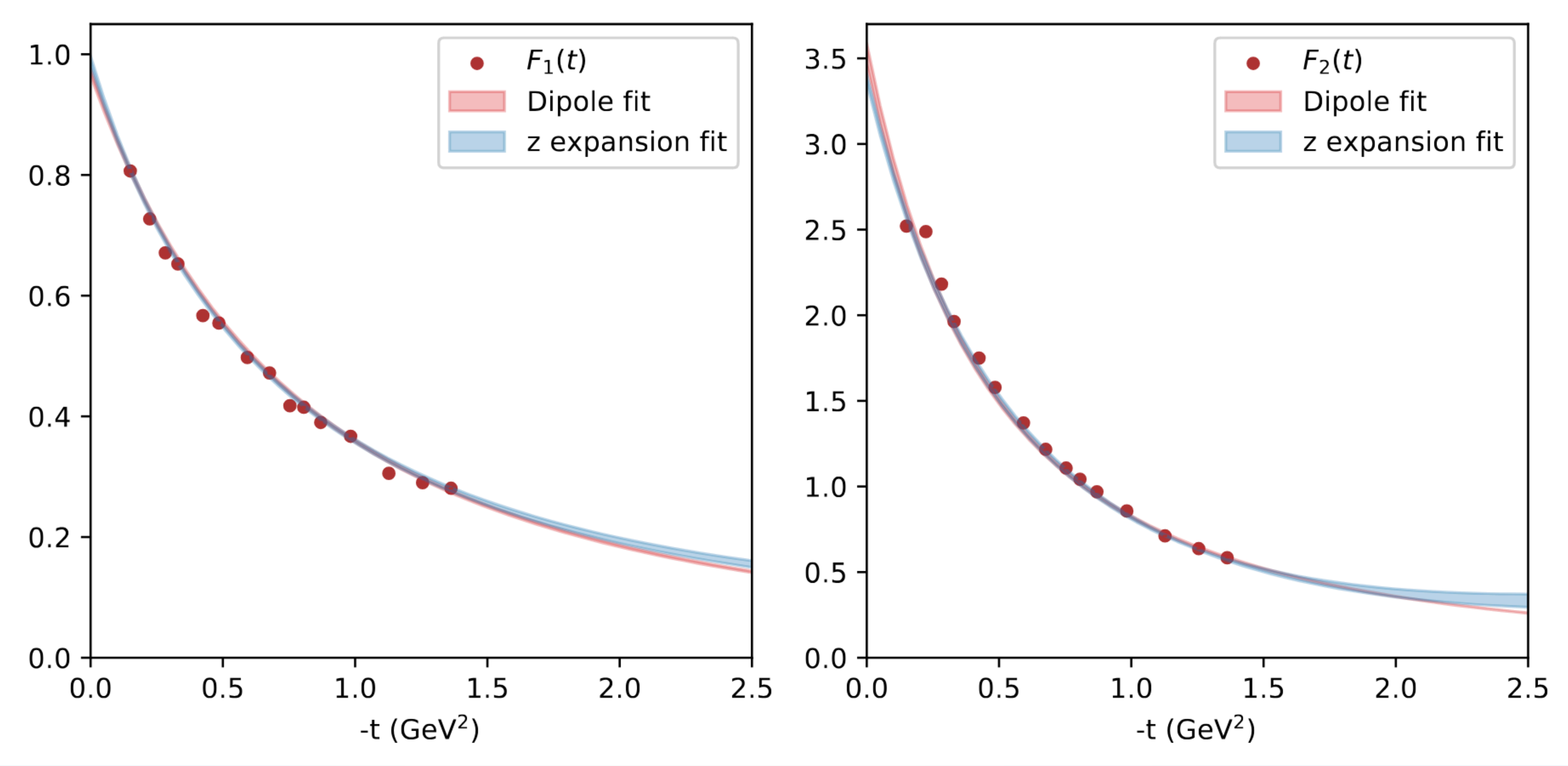
Planning to add the upcoming CLAS cross-section data.

Jefferson Lab

State of the project: Inputs from Lattice QCD

Electromagnetic form factors from Lattice QCD

- The obtained fit will be used as a constraint for the neural network.



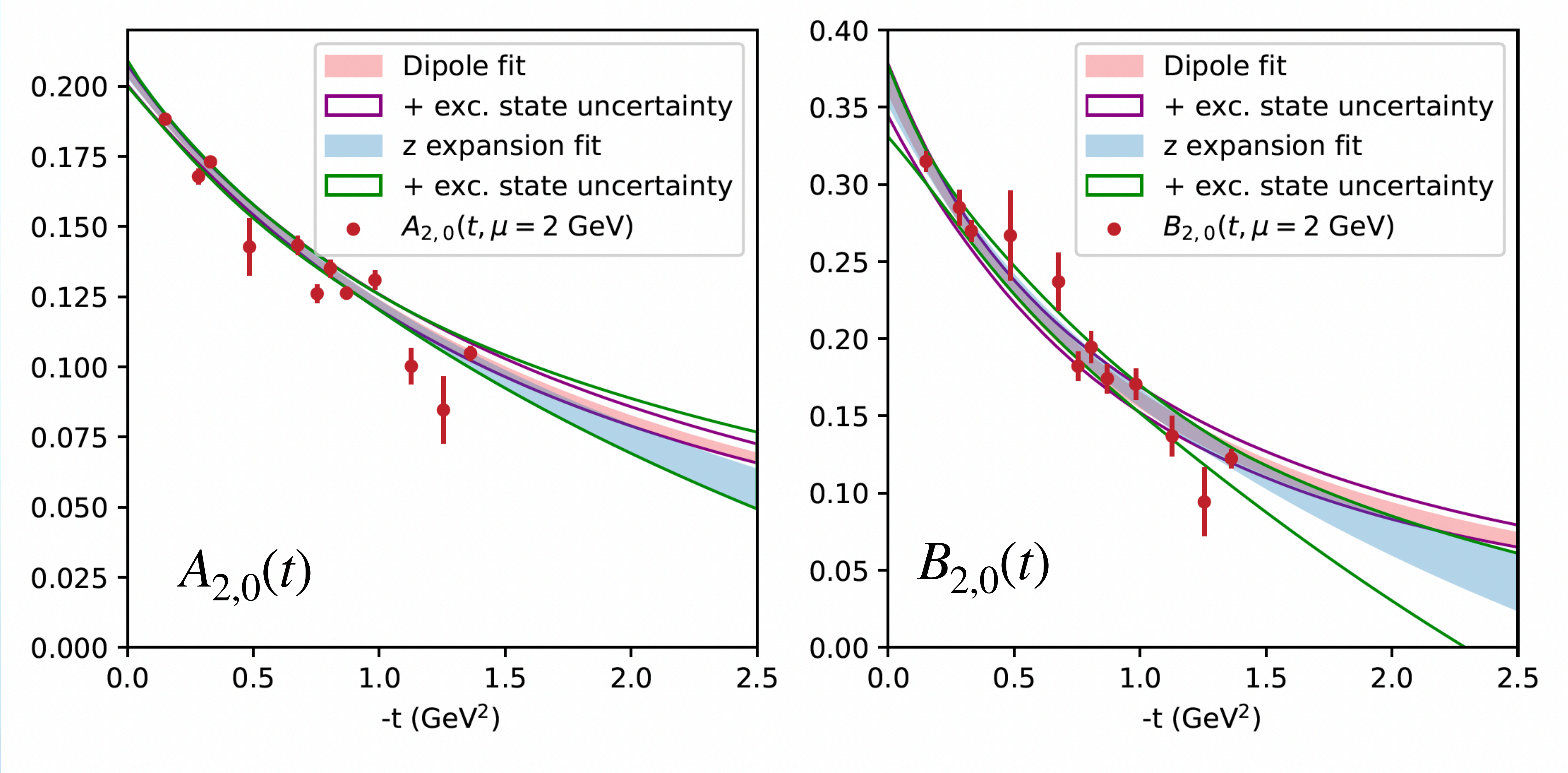
J. High Energy. Phys. 2024, 162 (2024)

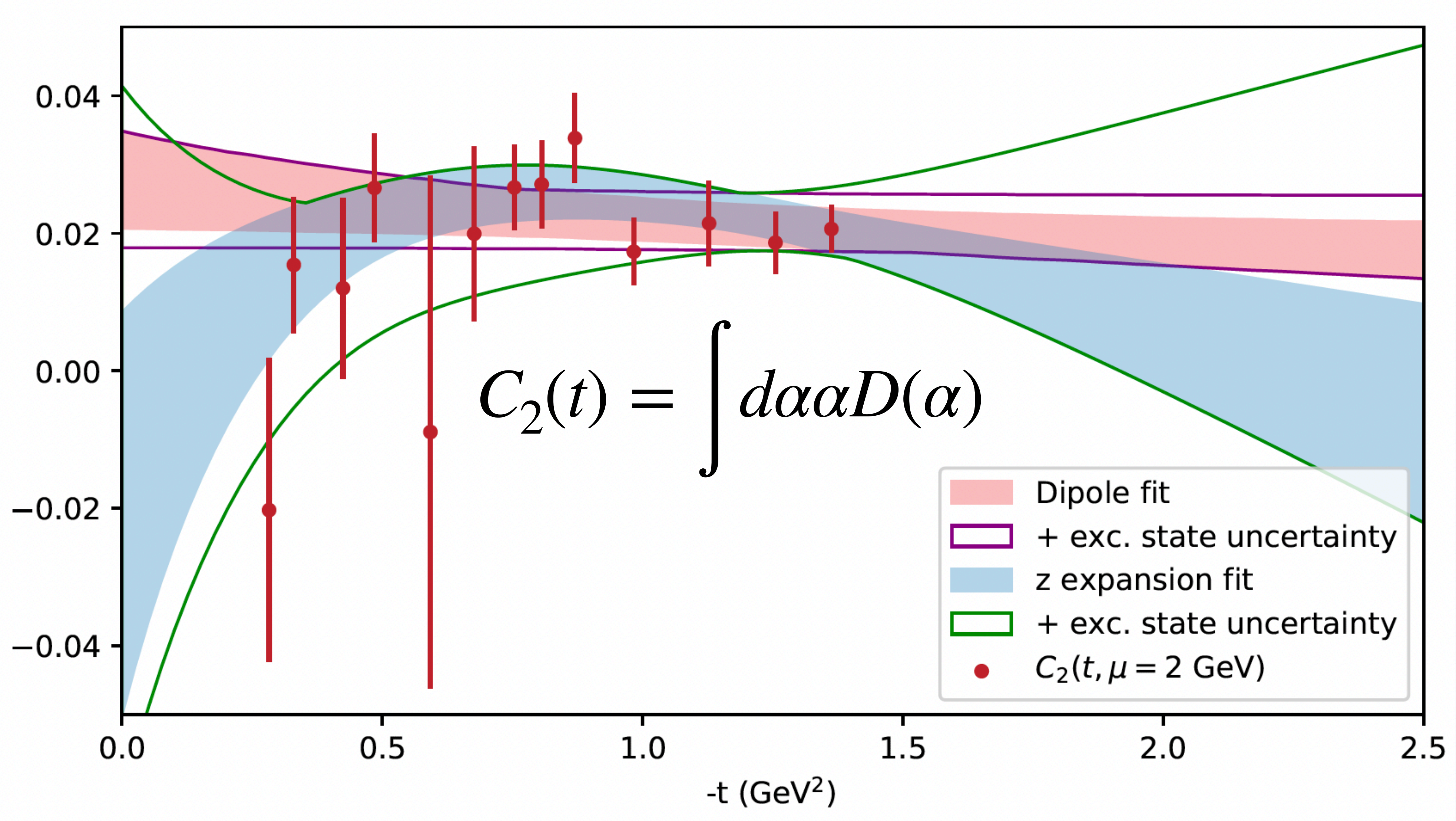
Jefferson Lab

State of the project: Inputs from Lattice QCD

Extraction of the isovector GFFs from Lattice QCD

- The non-singlet "$u-d$" combination of the GFF, which cannot be probed through DVCS, has been calculated in Lattice across a wide range of t.
- Needs of the isoscalar GFFs.





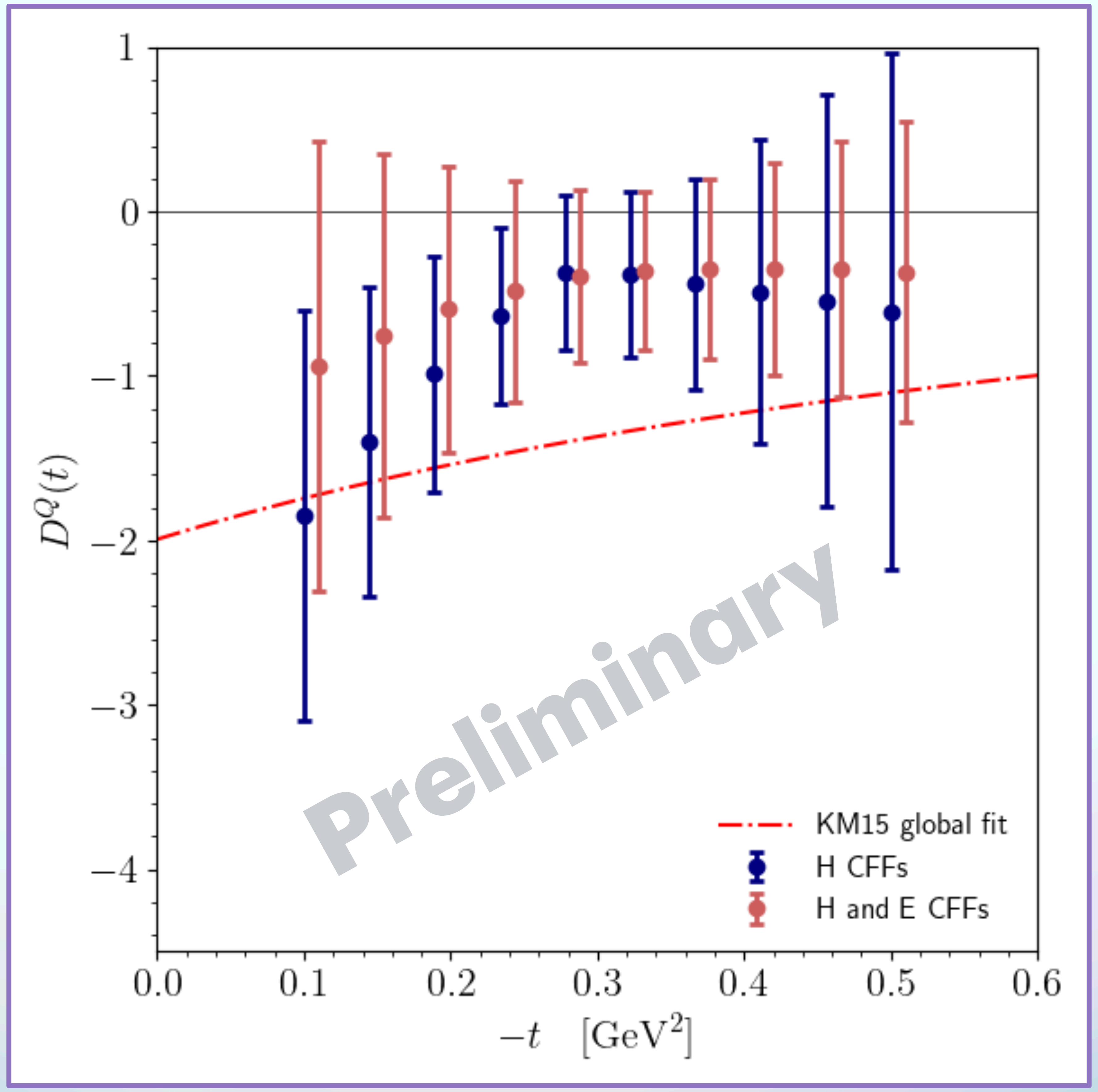
J. High Energy. Phys. 2024, 162 (2024)

Jefferson Lab

State of the project: D-term

Finally, the subtraction constant obtained from the NN leads to the following D-term values:

$$D^Q(t) = \frac{18}{25}\Delta(t)$$



- More data needs to be added to reduce uncertainties as well as a bigger ensemble for the NNets

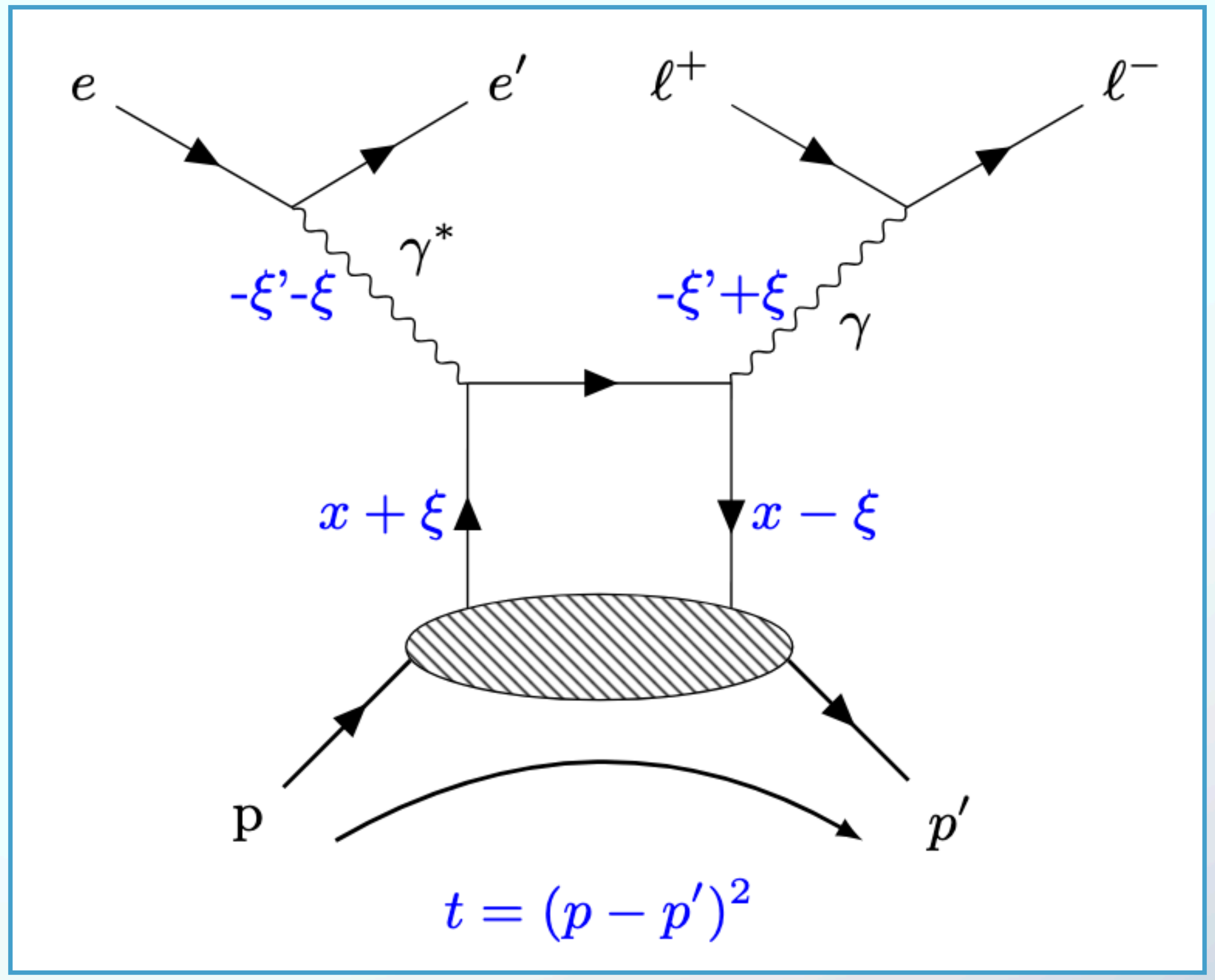
Deliverables and future directions

Double DVCS (DDVCS)

PoS DIS2024 (2025) 239

Another privilege process for GPDs extraction is the Double DVCS

- Possibility of directly measuring GPDs for $x \neq \xi$
- Providing a potential solution to deconvolution challenges.
- Better estimations of the moments of GPDs.



Monday's talk by Victor Martínez-Fernández

K. Deja, et. al, *Acta Phys.Polon.Supp. 16 (2023) 7, 7-A24*

SoLID new proposal

Deliverables and future directions

- Completed implementation of DR with PyTorch into Gepard
- A tool to fit all existing data using NN to extract GFFs capable of handling both experimental and LQCD data.
- Understand physical constraints from LQCD and implement them in the network architecture.
- Preliminary CFFs and D-term results for all CLAS data have been extracted.

Future directions on Lattice QCD computation:

- Calculate the isoscalar GFFs using Lattice QCD. I.J. High Energy. Phys. 2024, 162 (2024)
- Implement physical constraints from LQCD in the network architecture.

Future directions on the fitting framework:

- Implement DDVCS process into Gepard.
- Optimization of NNets of Gepard, for instance, implementation of the possibility for use GPU device.
- Implement NN to PARTONS which is a C++ framework for the phenomenology of the 3D structure of the nucleon (TMDs & GPDs).
- Global analysis by steps: add HallA data and the rest of DVCS data.

Thank you all for your attention

Other details

Data used thus far:

- First, we used 2008 beam spin asymmetry and 2015 cross-section CLAS data to reproduce previous results.
- Therefore new 2023 BSA CLAS12 data at 10.6 and 10.2 GeV (Phys. Rev. Lett. 130, 211902) was implemented to raise kinematic restrictions, and then the rest of CLAS data was added.
- Planning to use new 2022 Hall A data (Phys. Rev. Lett. 128, 252002)
- Planning to use upcoming cross-section CLAS12 data.

Neural Network construction in Gepard:

- Fits are performed in harmonic space, therefore harmonic analysis (Fourier transform) of the data was previously necessary.
- Trained 10 neural nets per model with an ensemble of at least 15 trainings.
- The architecture of nets parametrizing imaginary part of CFF is:
  - 2 input neurons (for x_B and t)
  - 17 and 25 neurons in hidden layers for reproducing previous results.
  - 20 and 25 neurons in hidden layers for the new CLAS data
  - 1 neuron in the output layer for each required imaginary part of CFF as well as the subtraction constant.